日本国特許庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日
Date of Application: ２００４年　２月　３日

出願番号
Application Number: 特願２００４－０２６３５６

[ST. 10/C]: [ＪＰ２００４－０２６３５６]

出願人
Applicant(s): 株式会社日立製作所

２００４年　２月１９日

特許庁長官
Commissioner,
Japan Patent Office

今井康夫

出証番号　出証特２００４－３０１１３１１

【書類名】　特許願
【整理番号】　NT03P0796
【提出日】　平成16年 2月 3日
【あて先】　特許庁長官　殿
【国際特許分類】　G06F 12/00
【発明者】
　【住所又は居所】　神奈川県川崎市麻生区王禅寺１０９９番地　株式会社日立製作所　システム開発研究所内
　【氏名】　江口　賢哲
【発明者】
　【住所又は居所】　神奈川県川崎市麻生区王禅寺１０９９番地　株式会社日立製作所　システム開発研究所内
　【氏名】　山本　康友
【発明者】
　【住所又は居所】　神奈川県川崎市麻生区王禅寺１０９９番地　株式会社日立製作所　システム開発研究所内
　【氏名】　荒川　敬史
【発明者】
　【住所又は居所】　神奈川県小田原市中里３２２番地２号　株式会社日立製作所　ＲＡＩＤシステム事業部内
　【氏名】　平川　裕介
【特許出願人】
　【識別番号】　000005108
　【氏名又は名称】　株式会社日立製作所
【代理人】
　【識別番号】　100068504
　【弁理士】
　【氏名又は名称】　小川　勝男
　【電話番号】　03-3537-1621
【選任した代理人】
　【識別番号】　100086656
　【弁理士】
　【氏名又は名称】　田中　恭助
　【電話番号】　03-3537-1621
【選任した代理人】
　【識別番号】　100094352
　【弁理士】
　【氏名又は名称】　佐々木　孝
　【電話番号】　03-3537-1621
【手数料の表示】
　【予納台帳番号】　081423
　【納付金額】　21,000円
【提出物件の目録】
　【物件名】　特許請求の範囲　1
　【物件名】　明細書　1
　【物件名】　図面　1
　【物件名】　要約書　1

【書類名】特許請求の範囲
【請求項１】
ホストと通信路を介して接続される第１の記憶装置システム及び第２の記憶装置システムを含み、第２の記憶装置システムは該第１の記憶装置システムから送られるデータのコピーを記憶するストレージサブシステムにおける、該第２の記憶装置システムに記憶されたデータを復元するデータの復元方法であって、
第１の記憶装置システムは、該ホストからの入出力要求を処理し、かつ第２の記憶装置システムに対して入出力処理の結果、更新されたデータを送信し、
第２の記憶装置システムは、第１の記憶装置システムからの受信したデータを更新ログデータとして保存し、
ホストはアプリケーションの状態確定するコマンドをデータとして第１の記憶装置システムに送信し、第１の記憶装置システムは該データを第２の記憶装置システムに送信し、
かつ該ホストと第２の記憶装置システムは、該コマンドに対応した識別子を双方で保持し、該識別子とログデータとを関連付けることにより、該ホストが任意の時点で該識別子を指示することによって第２の記憶装置システムで任意の時点のデータを復元することを特徴とするデータの復元方法。
【請求項２】
前記ホストは、リモートサイトにある該第２の記憶装置システムに識別子の入出力指示を発行することを特徴とする請求項１のデータの復元方法。
【請求項３】
リモートサイトにある該第２の記憶装置システムは、ホストの識別子の入出力指示を受領し、データの更新ログと該識別子を関連つけて記憶装置に記憶すること特徴とする請求項１のデータの復元方法。
【請求項４】
第２の記憶装置システムに記憶されているデータを復元する場合、該ホストから送信され、受信した識別子と一致する識別子を検索し、対象とする識別子を捜したらコピー先の記憶装置に格納されたデータと、一致した識別子と関連付けられたログデータより前に記録されたログデータの内容を用いて、コピー元の記憶装置にデータを復元することを特徴とする請求項１のデータの復元方法。
【請求項５】
該第１の記憶装置システムは、ホストからログデータの取得開始コマンド、及び該記憶装置のペア状態を中断するコマンドを受領すると、ペア関係にある第２の記憶装置システムにある記憶装置を確認して、ペアの状態をサスペンドすることを特徴とする請求項１乃至４のいずれかのデータの復元方法。
【請求項６】
該第２の記憶装置システムは、該ホストで発行されたマークコマンドを受領すると、ログデータを取得するする記憶装置を確認し、かつ
取得されたログデータに対してマークＩＤ及びタイマ値を含むマークデータの対応を設定することを特徴とする請求項１乃至５のいずれかのデータの復元方法。
【請求項７】
ホストとネットワークによって接続された記憶装置を備えた記憶装置システムを含む計算機システムにおける該記憶装置に記憶されるデータの処理方法において、
ホストは、記憶装置システムに対して記憶装置に格納されているデータのコピーの作成、保存を要求するステップと、ホストの処理によるデータの更新部分の記録を要求するステップと、計算機システムのある時点の状態を識別する識別情報を記憶装置システムに送信するステップと、
該記憶装置システムは、ホストの要求に応答して、該記憶装置のデータのコピーを作成して保存するステップと、該記憶装置の内容が更新されたときに、更新前後のデータ及び更新場所を示す情報をログデータとして保存するステップと、ホストより送信される識別情報を保持するステップと、該ログデータと識別情報を関連付けるステップと、

を有することを特徴とするデータ処理方法。
【請求項８】
該記憶装置に記憶された内容をある時点の状態に復旧する場合、ホストは状態識別情報を指定してデータの復旧要求を記憶装置システムに送信するステップと、
該記憶装置システムは、受信された該識別情報を判別し、該データのコピーとログデータを用いてデータをリストアするステップと、
を更に有すること特徴とする請求項７記載のデータの復元方法。
【請求項９】
該ホスト及び記憶装置システムで共有される該識別情報を、該記憶装置システムで識別情報と更新履歴を関連付けて管理し、該ホストからの指示に応じて、特定の識別情報で示される更新履歴までの該記憶装置に格納されたデータを復元することを特徴とする請求項８記載のデータの復元方法。
【請求項１０】
前記ログデータは少なくとも、特定の識別情報であるか否かを示すマークフラグのためのエントリと、ログの識別情報のためのエントリと、該ログを取得した時間を示すタイマ値ためのエントリと、データ領域となるエントリとを含み、
該マークフラグが特定の識別情報を示す場合、該データ領域のエントリは、少なくともマークの識別情報のエントリと、マークを付与した時間を示すタイマ値のためのエントリを表すエントリを規定することを特徴とする請求項８又は９記載のデータの復元方法。
【請求項１１】
該ホストは、プログラム実行に際してファイルをクローズする時、又はファイルをセーブする時に、該ログデータと識別情報を関連付けるための特定のコマンドを生成して該記憶装置システムに送信し、該記憶装置システムで該コマンドを実行して、該マークフラグを特定の状態にセットし、かつ前記エントリにマークの識別情報とタイマ値を格納したログデータを記憶装置に保存することを特徴とする請求項１０記載のデータの復元方法。
【請求項１２】
ホストと通信路を介して接続される第１の記憶装置システム及び第２の記憶装置システムを含み、該第１の記憶装置システムに記憶するデータのコピーを第２の記憶装置システムに記憶するストレージサブシステムにおいて、
該第１の記憶装置システムは；
複数の論理記憶装置を有する記憶装置と、
該記憶装置に入力又は出力されるデータを一時的に格納するキャッシュメモリと、
少なくとも該論理記憶装置に関する管理情報、該第１の記憶装置システムと第２の記憶装置システムのペア関係の構成を規定する管理情報、および該ホストからのコマンドを処理するプログラムを格納するメモリと、
該プログラムを実行するプロセッサと、を有し、
該第２の記憶装置システムは；
複数の論理記憶装置を有し、その内ある論理記憶装置はペアを構成する該第1の記憶装置システムの論理記憶装置に記憶されるデータのコピーを格納し、他のある論理記憶装置は該第１の記憶装置システムで生成されたログデータを記憶するために割当てられる記憶装置と、
該記憶装置に入力又は出力されるデータを一時的に格納するキャッシュメモリと、
少なくとも該論理記憶装置に関する管理情報、該第１の記憶装置システムと第２の記憶装置システムのペア関係の構成を規定する管理情報、ログの管理情報、および該ホストからのコマンドを処理するプログラムを格納するメモリと、
該プログラムを実行するプロセッサと、を有し、
該第1の記憶装置システムのある論理記憶装置の内容が更新された場合、更新されたデータ及び更新場所を示す情報をログデータとして該第２の記憶装置システムに送信して、該論理記憶装置に格納し、かつ該該該ホストから送信された識別情報と該ログデータを対応付ける管理情報を該メモリに格納することを有することを特徴とするストレージサブシス

テム。
【請求項１３】
該第２の記憶装置システムにおける該論理記憶装置に格納されたログデータをある時点の状態に復旧する場合、該第２の記憶装置システムは、ホストから送信されたデータの復旧要求を受信して、該識別情報に関して該メモリに格納された前記管理情報を参照して該論理記憶装置に格納されたログデータをリストアすることを特徴とする請求項１２記載のデータストレージサブシステム。
【請求項１４】
該第１の記憶装置システムの該メモリ及び該第２の記憶装置システムの該メモリは、
該ホストが識別できる該記憶装置のアドレスと、該記憶装置内の論理アドレスの対応を登録する記憶装置管理情報テーブルと、
該記憶装置内の論理アドレスと、該論理記憶装置が配置されているＲＡＩＤグループに関するアドレスと、該ＲＡＩＤグループを形成するディスクに関するアドレスの対応を登録する記憶装置管理情報テーブルと、
論理記憶装置の番号に対応してボリューム構成情報を登録するボリューム構成情報テーブルと、
ペアを構成する該第１の記憶装置システム内の論理記憶装置の番号と該第２の記憶装置システム内の論理記憶装置番号の対応を登録するペア管理情報テーブルと、
を有することを特徴とする請求項１２又は１３記載のデータストレージサブシステム。
【請求項１５】
少なくとも該第２の記憶装置システムは、
ログボリュームグループごとにログボリュームグループ内の論理記憶装置に関する管理情報を登録するログボリュームグループ情報テーブルと、
ログデータを該論理記憶装置に格納されるときに付与されるログＩＤと、ログデータが取得された時のタイマ値を対応付けて管理情報として登録するログＩＤ管理情報テーブルと、
ログデータが該論理記憶装置に格納されるときに付与される、ホストから送信された識別情報と、ログデータが格納される時のタイマ値を対応付けて管理情報として登録するマークＩＤ管理情報テーブルと、
を有することを特徴とする請求項１４記載のデータストレージサブシステム。
【請求項１６】
前記第１の記憶装置システムにおいて、
前記コマンド処理プログラムは、該ホストから送信されるコマンドを処理し、該コマンドがマークコマンドの場合には、ログデータを作成して、該識別情報を付与するための処理を行い、
Ｉ／Ｏ処理コマンドの場合には、該キャッシュメモリのヒット又はミスヒットを判定し、ライトデータを該キャッシュメモリに書き込み、又は該キャッシュメモリからリードデータを読み出す処理を行うことを特徴とする請求項１２乃至１５のいずれか記載のデータストレージサブシステム。
【請求項１７】
該第２の記憶装置システムは、該ホストで発行されたマークコマンドを受領すると、ログデータを取得するする記憶装置を確認し、かつ取得されたログデータに対してマークＩＤ及びタイマ値を含むマークデータの対応を設定することを特徴とする請求項１２乃至１６のいずれか記載のデータストレージサブシステム。
【請求項１８】
ホストと通信路を介して接続される該第１の記憶装置システムに記憶されるデータのコピーを第２の記憶装置システムにおいて、
複数の論理記憶装置を有し、その内ある論理記憶装置はペアを構成する該第1の記憶装置システムの論理記憶装置に記憶されるデータのコピーを格納し、他のある論理記憶装置は該第１の記憶装置システムで生成されたログデータを記憶するために割当てられる記憶装

置と、
該記憶装置に入力又は出力されるデータを一時的に格納するキャッシュメモリと、
少なくとも該論理記憶装置に関する管理情報、該第１の記憶装置システムと第２の記憶装置システムのペア関係の構成を規定する管理情報、ログの管理情報、および該ホストからのコマンドを処理するプログラムを格納するメモリと、
該プログラムを実行するプロセッサと、を有し、
該第1の記憶装置システムのある論理記憶装置の内容が更新された場合、更新されたデータ及び更新場所を示す情報をログデータとして該第２の記憶装置システムに送信して、該論理記憶装置に格納し、かつ該該該ホストから送信された識別情報と該ログデータを対応付ける管理情報を該メモリに格納することを有することを特徴とする記憶装置システム。
【請求項１９】
該論理記憶装置に格納されたログデータをある時点の状態に復旧する場合、該第２の記憶装置システムは、ホストから送信されたデータの復旧要求を受信して、該識別情報に関して該メモリに格納された前記管理情報を参照して該論理記憶装置に格納されたログデータをリストアすることを特徴とする請求項１８記載の記憶装置システム。
【請求項２０】
ホストと通信路を介して接続される第１の記憶装置システム及び第２の記憶装置システムを含み、該第１の記憶装置システムに記憶するデータのコピーを第２の記憶装置システムに記憶するストレージサブシステムにおいて、
該第１の記憶装置システムは；
複数の論理記憶装置を有する記憶装置と、
該記憶装置に入力又は出力されるデータを一時的に格納するキャッシュメモリと、
少なくとも該論理記憶装置に関する管理情報、該第１の記憶装置システムと第２の記憶装置システムのペア関係の構成を規定する管理情報、および該ホストからのコマンドを処理するプログラムを格納するメモリと、
該プログラムを実行するプロセッサと、を有し、
前記コマンド処理プログラムは、該ホストから送信されるコマンドを処理し、該コマンドがマークコマンドの場合には、ログデータを作成して、該識別情報を付与するための処理を行い、Ｉ／Ｏ処理コマンドの場合には、該キャッシュメモリのヒット又はミスヒットを判定し、ライトデータを該キャッシュメモリに書き込み、又は該キャッシュメモリからリードデータを読み出す処理を行い、
該第２の記憶装置システムは；
複数の論理記憶装置を有し、その内ある論理記憶装置はペアを構成する該第1の記憶装置システムの論理記憶装置に記憶されるデータのコピーを格納し、他のある論理記憶装置は該第１の記憶装置システムで生成されたログデータを記憶するために割当てられる記憶装置と、
該記憶装置に入力又は出力されるデータを一時的に格納するキャッシュメモリと、
少なくとも該論理記憶装置に関する管理情報、該第１の記憶装置システムと第２の記憶装置システムのペア関係の構成を規定する管理情報、ログの管理情報、および該ホストからのコマンドを処理するプログラムを格納するメモリと、
該プログラムを実行するプロセッサと、を有し、
該第1の記憶装置システムのある論理記憶装置の内容が更新された場合、更新されたデータ及び更新場所を示す情報をログデータとして該第２の記憶装置システムに送信して、該論理記憶装置に格納し、かつ該該該ホストから送信された識別情報と該ログデータを対応付ける管理情報を該メモリに格納し、
該第１の記憶装置システムは、該ホストからログデータの取得開始コマンド、及び該記憶装置のペア状態を中断するコマンドを受領すると、ペア関係にある第２の記憶装置システムにある記憶装置を確認して、ペアの状態をサスペンドし、
該第２の記憶装置システムにおける該論理記憶装置に格納されたログデータをある時点の状態に復旧する場合、該第２の記憶装置システムは、ホストから送信されたデータの復旧

要求を受信して、該識別情報に関して該メモリに格納された前記管理情報を参照して該論理記憶装置に格納されたログデータをリストアすることを特徴とするデータストレージサブシステム。

【書類名】明細書
【発明の名称】ストレージサブシステム
【技術分野】
【０００１】
本発明は、ストレージサブシステムに係り、特にライトＩ／Ｏログを取得するストレージサブシステムの遠隔コピー、及び災害時におけるログの復旧方法に関する。
【背景技術】
【０００２】
計算機と記憶装置システムとをネットワークにより接続し、計算機で処理されるデータを、ネットワークを介して送受信して記憶装置システムに格納する計算機システムが実用化されている。計算機システムで実行されるオンライン処理やバッチ処理において、プログラムのバグや記憶装置システムの障害などによってこれらの処理が異常終了し、記憶装置システムに格納されたデータに矛盾が生じる事態になることがある。また、人為的ミスによって記憶装置システムに格納されたデータが消去されてしまうこともある。
【０００３】
このような状態になった計算機システムのデータを回復させるために、データの矛盾を解消して途中で止まった処理を再開させたり、あるいは途中で止まった処理をもう一度実行し直したりするための技術の一つとして、データのバックアップとリストアによるデータ回復技術が知られている。
【０００４】
バックアップおよびリストアに関する従来技術として、例えば米国特許番号５，２６３，１５４号公報（特許文献１）には、ユーザが指定した時点における記憶装置システムに格納されたデータを、記憶装置システムに接続されたホストからのデータの入出力（以下、「Ｉ／Ｏ」）を止めることなく磁気テープに複製し（以下、「データのバックアップ」）、その複製されたデータ（以下、「バックアップデータ」）を用いてデータの回復（以下、「リストア」）する技術が開示されている。
【０００５】
また、特開２００１－２１６１８５号公報（特許文献２）には、データのリストアにかかる時間を短縮するために、データのバックアップが実行された後、データが更新された場所における情報を差分情報として保持し、記憶装置システムに格納されたデータをバックアップデータでリストアする際に、バックアップデータのうち、差分情報で示されるデータの部分のみをデータのリストアに用いる技術が開示されている。
【０００６】
また、米国特許番号５，５４４，３４７号公報（特許文献３）、米国特許番号５，７４２，７９２号公報（特許文献４）には、リモートにある記憶装置システムで、ホストに独立に、データをコピーするという技術が開示されている。この技術によれば、ある業務サイトにてホストによって作成された、記憶装置システム内の業務ボリュームのコピーをリモートサイトの記憶装置システム内のボリュームに作成することができる。そのため、業務サイトが天災やテロなどの被災に会い、計算機システムが障害を起こし、業務が継続できなくなった場合を考慮して、リモートサイトにおいて業務の復旧のためのデータが用意することができる。
【０００７】
【特許文献１】米国特許番号５，２６３，１５４号公報
【０００８】
【特許文献２】特開２００１－２１６１８５号公報
【特許文献３】米国特許番号５，５４４，３４７号公報
【特許文献４】米国特許番号５，７４２，７９２号公報
【発明の開示】
【発明が解決しようとする課題】
【０００９】

特許文献１に記載されたリストア処理では、磁気テープからバックアップデータを読み出す際、バックアップデータを取得した時点から更新されていない部分（即ち、記憶装置システムのデータと磁気テープのデータの内容が一致している部分）も磁気テープから読み出され、記憶装置システムに書き込まれる。このようなデータの転送は、無駄が多く、リストアのための時間が長くなる。

【００１０】

特許文献２に開示されている技術では、特許文献１の技術に比べ、重複したデータの読み出しが発生しない分、リストアに係る時間は少なくなる。しかし、双方の技術をもってしても、データのバックアップの後から記憶装置システムが故障するまでの間に更新されたデータについては、データのリストアを行うことができない。データのバックアップ後に更新されたデータまでリストアしようとすると、そのデータの更新の内容等をホスト側がログ等で管理する必要がある。このためホストへの負荷が大きく、またその処理に長い時間がかかる。

【００１１】

また、特許文献３や特許文献４には、業務サイトが被災し、計算機システムが障害を起こして業務が継続できなくなった場合を想定して、リモートサイトで復旧のためのデータを作成する技術が開示されているが、災害中の書き込みにより生じる不正な書き込みに関するリモートサイトでのコピーの危険性に対する対策案については考慮されていない。

【００１２】

本発明の目的は、障害発生前までの任意の時点におけるデータのリストア処理を高速に行う計算機システムを提供することにある。

【００１３】

本発明の目的は、記憶装置システムに格納されたデータを復旧する場合に、ホストに負担をかけず、リモートサイトでデータをリストアする記憶装置システムを提供することにある。

【課題を解決するための手段】

【００１４】

本発明は、好ましくは、ホストと通信路を介して接続される第１の記憶装置システム及び第２の記憶装置システムを含み、第２の記憶装置システムは第１の記憶装置システムから送られるデータのコピーを記憶するストレージサブシステムにおいて実現される。

このストレージサブシステムにおいて、第１の記憶装置システムは、ホストからの入出力要求を処理し、かつ第２の記憶装置システムに対して入出力処理の結果、更新されたデータを送信し、第２の記憶装置システムは、第１の記憶装置システムからの受信したデータを更新ログデータとして保存し、ホストはアプリケーションの状態確定するコマンドをデータとして第１の記憶装置システムに送信し、第１の記憶装置システムはそのデータを第２の記憶装置システムに送信し、かつホストと第２の記憶装置システムは、コマンドに対応した識別子を双方で保持し、その識別子とログデータとを関連付ける。データを復元する場合に、ホストは任意の時点で識別子を指示することによって第２の記憶装置システムで任意の時点のデータを復元する。

好ましい例では、ホストは、リモートサイトにある第２の記憶装置システムに状態識別情報の入出力指示を発行する。リモートサイトにある第２の記憶装置システムは、ホストの状態識別情報の入出力指示を受領し、データの更新ログと識別情報を関連つけて記憶装置に記憶する。

【発明の効果】

【００１５】

本発明によれば、記憶装置システムに格納されたデータを復旧する場合に、ホストに負担をかけず、リモートサイトで短時間にデータを所定の状態までリストアすることができる。

【発明を実施するための最良の形態】

【００１６】

以下、図面を用いて、本発明の一実施形態について説明する。
図１は、一実施形態による計算機システムのブロック図を示す。
この計算機システムは、第１のホスト計算機１（以下、ホスト１という）と、第２のホスト計算機２（以下、ホスト２という）と、ホスト１に通信路５を介して接続される第１の記憶装置システム３（以下単に記憶装置システム３という）と、ホスト２に通信路７を介して接続されると共に、記憶装置システム３と通信路６を介して接続される第２の記憶装置システム４（以下単に記憶装置システム４という）と、記憶装置システム３に通信路１１を介して接続される第１の記憶装置システム管理装置８と、記憶装置システム４に通信路１２を介して接続される第２の記憶装置システム管理装置（以下、管理装置という）９と、これらの構成装置と通信路１０を介して接続される計算機システム管理装置１３を含んで構成される。記憶装置システム３，４は、例えばディスク装置或いはディスクアレイのような記憶装置２１，２２，２３及びそれらの制御装置を有して構成される。ここで第２系の側はリモートサイトとして位置付けされる。

【００１７】

ホスト１及びホスト２は、パーソナルコンピュータ、ワークステーション、メインフレーム等の計算機である。ホスト１では、その計算機の種類に応じたオペレーティングシステム（ＯＳ）や、様々な業務や用途に対応したアプリケーションプログラム（ＡＰ）例えばデータベース（ＤＢ）プログラムのような色々なプログラムが実行される。なおこの例では、簡単のため、ホスト１とホスト２とは１つずつ示されているが、通信路５または通信路７に接続されるホスト１またはホスト２は複数であってよい。
ホスト１は、計算機システムにおける所定の処理を実行する。即ち、ホスト１は情報処理に係る入出力処理に必要なコマンド及びデータを、記憶装置システム３との間で通信路５を用いて通信し、ホスト１で作成、変更したデータを記憶装置システム３に対してデータのライト要求を行い、また計算機処理に必要なデータのリード要求を行う。

【００１８】

記憶装置システム３は、通信路５を介して送信されるコマンドやデータを受信し、所定の処理を行い、ホスト１に対する所定の応答を実行する。記憶装置システム３と記憶装置システム４は、通信路６を介してそれらの間でコマンド及びデータの通信を行う。
記憶装置システム４の構成も記憶装置システム３と実質的に同じであるが、図１では複数の記憶装置２２、２３を有する点が相違する。詳しくは後述するが、２つある記憶装置の内一方の記憶装置２２は記憶装置２１のミラー用であり、他方の記憶装置２３はログデータ記憶用に使用される。

【００１９】

管理装置８は、記憶装置システム３の設定を行う他、記憶装置システム３の障害、保守、構成、性能情報等の管理を行うための計算機である。同様に、管理装置９は、記憶装置システム４の設定を行うと共に、記憶装置システム３の障害、保守、構成、性能情報等の管理を行うための計算機である。例えば、計算機システムの管理者が、記憶装置システム３及び記憶装置システム４に、例えばボリュームのような論理的な記憶装置を設定する場合、データをバックアップするための記憶領域を設定する場合、又はデータを複製する際の記憶領域の対を設定する場合に、これらの管理装置８、９が用いられる。
システム管理者は、記憶装置システム３の保守・管理、記憶装置システム３が有する物理記憶装置の設定、及び記憶装置システム３と接続されるホスト１の設定等を行う場合に、管理装置８に設定したい内容を入力する。システム管理者が管理装置８に入力した内容は、通信路１１、１０を介して記憶装置システム３及びホスト１に送信される。
計算機システム管理装置１３は、計算機システム全体の保守、管理を行うためのものであり、通常システムエンジニアにより操作され利用される。

【００２０】

通信路５及び７は、ホスト１、ホスト２が夫々記憶装置システム３、記憶装置システム４へＩ／Ｏの処理要求等を伝送するために使用される。通信路６は、記憶装置システム３と記憶装置システム４との間のＩ／Ｏの処理要求等を伝送するために使用される。通信路

５，６，７には、光ケーブルや銅線等が用いられる。これらの通信路５，６，７で使用される通信プロトコルとしては、例えばイーサネット（登録商標）、ＦＤＤＩ、ファイバチャネル、ＳＣＳＩ、Ｉｎｆｉｎｉｂａｎｄ、ＴＣＰ／ＩＰ、ｉＳＣＳＩなどがある。
通信路１１及び通信路１２は、記憶装置システム３や記憶装置システム４が、自身の障害、保守、構成、性能等の管理情報を管理装置８、９との間で転送するために使用される。
通信路１０は、ホスト１、２が管理装置８、９から管理情報を取得する場合にコマンドを転送したり、及びに記憶装置システム３、４の障害、保守、構成、性能等の管理情報を管理装置８、９からホスト１に送信したり使用される。通信路１０，１１，１２で使用されるケーブル及び通信プロトコルは、通信路５，６，７と同じでもよいし、異なっていてもよい。

【００２１】

図２は、記憶装置システム３、４の構成を示す図である。
記憶装置システム３は、記憶装置システム制御装置１０１と記憶装置２１から構成され、記憶装置２１の記憶領域はホスト１からの入出力のために提供される。即ち記憶装置システム３は、ホスト１が使用するデータやプログラムを格納し、ホスト１のＩ／Ｏ処理要求を受信し、Ｉ／Ｏ処理要求に対応した処理を行い、その結果を所定のホスト１に送信する。

【００２２】

記憶制御装置１０１は、記憶装置２１と接続される記憶装置アダプタ１０８、所定のプログラムを実行するプロセッサ１０５、プロセッサ１０５で実行されるプログラム、プログラムが動作する上で必要な情報、記憶装置システム３の設定情報及び構成情報等が格納される不揮発性制御メモリ１０７、この記憶装置システム３を通信路１１と接続するためのネットワークアダプタ１０２、及び通信路５と接続するためのホストアダプタ１０３、通信路６と接続するためのリモートＩ／Ｏアダプタ１０４を有し、ホスト１からのＩ／Ｏ要求の処理及び記憶装置２１の制御を行う。

【００２３】

記憶装置２１は、ボリユームと称する複数の論理的な記憶装置１０９から構成され、夫々の記憶装置１０９にはユーザ用のデータが格納される。記憶装置２１は、その記憶領域に障害が生じてもデータが損失しないように、冗長性を持つＲＡＩＤ(Redundant Arrays of Inexpensive Disks)構成とするのが好ましい。記憶装置２１の記憶媒体としては、例えば電気的に不揮発な記憶媒体である磁気ディスクや不揮発性半導体メモリで構成されるシリコンディスク、光ディスク、光磁気ディスク又はハードディスク等の媒体が使用される。
尚、記憶制御装置１０１は、記憶装置システム３に複数存在しても良い。また記憶装置システム３の冗長性を確保するために、好ましくは、記憶制御装置１０１内の各構成要素への電源供給のための回路、記憶制御装置１０１内の各構成要素間でのデータの送受信を行う回路、キャッシュメモリ１０６、不揮発性制御メモリ１０７、記憶装置アダプタ１０８等を、それぞれ２重化された冗長構成とするのがよい。

【００２４】

不揮発性メモリ１０７は、記憶装置システム３内の処理を制御するために、コマンド処理処理プログラム１１１、リモートＩ／Ｏ処理プログラム１１２、コピーマネージャプログラム１１３、ペア管理プログラム１１４、ネットワーク処理プログラム１１９、記憶装置Ｉ／Ｏ処理プログラム１２１、システム構成管理プログラム１２２の各プログラムを格納する。不揮発性メモリ１０７は、またＩ／Ｏ処理のためにキャッシュメモリ１０６のある領域へアクセスを行うための排他処理用ビットや記憶装置１０９とキャッシュメモリ１０７との対応関係を示す情報を格納する。尚、不揮発性の制御メモリ１０７は、記憶制御装置１０１やプロセッサ１０５が複数在る場合には、それらで共有するようにしてもよい。
キャッシュメモリ１０６は、ホスト１から記憶装置システム３に転送されるデータ又は記憶装置システム３からホスト１へ転送されるデータを一時的に格納する。

【００２５】
ローカルエリアネットワーク１２０は、記憶制御装置１０１、キャッシュメモリ１０６、及び記憶装置１０９を相互に接続する。ローカルエリアネットワーク１２０は、例えば共有バス型の構成でもよいし、スター型等のネットワーク構成でもよい。
【００２６】
記憶制御装置１０１は、不揮発性メモリ１０７に格納されたプログラムをプロセッサ１０４で実行することで、以下に説明する処理を制御する。
コマンド処理プログラム１１１は、ホスト１からのＩ／Ｏ処理要求をホストアダプタ１０２で受信すると、受信したＩ／Ｏ処理要求の内容を解析する。解析の結果に従って、Ｉ／Ｏ処理要求の内容がデータの読み出しＩ／Ｏ（以下「リードＩ／Ｏ」）要求やデータの書き込みＩ／Ｏ（以下「ライトＩ／Ｏ」）処理要求等を実行する。ライトＩ／Ｏ処理要求の場合、記憶制御装置１０１は、ホスト１からのライトＩ／Ｏ処理要求に対する応答処理、例えば実際にホスト1から転送されるデータを受領できる状態にあるかどうかの応答処理を行い、更に転送されてくる更新用のデータ（以下「ライトデータ」）をキャッシュメモリ１０６又は記憶装置１０９の指定された場所に書き込む。リードＩ／Ｏ処理要求の場合、記憶制御装置１０１は、リードＩ／Ｏ処理要求に対応するデータ（以下、「リードデータ」）を、キャッシュメモリ１０６もしくは記憶装置１０９の指定の場所から読み出してホスト１に転送する。またその他の処理、例えばホストとの入出力Ｉ／ＦがＳＣＳＩで、ＳＣＳＩのＩｎｑｕｉｒｙコマンド（デバイスサーチを指示するコマンド）を受領した場合、ホストＩ／Ｏ処理プログラムは、コマンドで要求される処理内容に対応した動作の制御を行う。
【００２７】
システム構成管理プログラム１２２は、記憶制御装置１０１が記憶装置システム３を管理する際に実行するプログラムであり、記憶装置管理情報１１５、１１６の作成、設定、変更、削除等を行う。記憶制御装置１０１は、システム構成管理プログラム１２２を実行することによって、管理装置８から入力される記憶装置１０９の定義やバックアップ／スナップショット用のペア対象となる記憶装置１０９の設定、ＬＯＧ格納対象情報の登録等を行う。ここで、記憶装置管理情報１１５、１１６は、記憶装置システム内のアドレスとホスト記憶装置に対して入出力を行うためのアドレスとの対応関係を示すマッピング情報、ペア１１８に関する情報を保持する。
【００２８】
記憶装置Ｉ／Ｏ処理プログラム１２１は、記憶制御装置１０１がキャッシュメモリ１０６又は記憶装置１０９に対してデータのリードライト処理を行う場合に実行されるプログラムである。リードライトＩ／Ｏ処理要求の実行に際しては、システム構成管理プログラム１２２を実行してアクセス先の記憶装置１０９の構成をチェックして、リードライトＩ／Ｏ処理要求が指定するリード又はライトデータが読み出され又は格納されるべき場所のアドレスがどの記憶装置１０９のアドレスに対応するかを計算し、その計算結果に基づいて、記憶装置１０９へのアクセスを行う。
【００２９】
ペア管理プログラム１１４は、マスタボリューム（以下、「Ｍ－ＶＯＬ」）とリモートボリューム（以下「Ｒ－ＶＯＬ」）のペアを管理するためのプログラムである。記憶制御装置１０１は、ペア管理プログラム１１４を実行することで、ホスト１からの指示に従って、ある記憶装置１０９（「Ｍ－ＶＯＬ」）及びＭ－ＶＯＬ１０９に格納されたデータの複製を格納するリモートの記憶装置システム４内にある記憶装置２２内の記憶装置（「Ｒ－ＶＯＬ」）１３９について、ペア作成（Ｐａｉｒ　Ｃｒｅａｔｅ）、ペア中断（Ｐａｉｒ　Ｓｕｓｐｅｎｄ）、ペア再開（Ｐａｉｒ　Ｒｅｓｙｎｃ）、ペア解除（Ｐａｉｒ　Ｄｅｌｅｔｅ）の処理を行う。
尚、１つのＭ－ＶＯＬに対して、複数のＲ－ＶＯＬを設定し、作成することもできる。また、Ｒ－ＶＯＬを新たなＭ－ＶＯＬとして、新たなＭ－ＶＯＬとペアになるＲ－ＶＯＬを設定、作成することもできる。

【００３０】
ペア管理情報１１８は、あるＭ－ＶＯＬ１０９がペア状態（Ｐａｉｒ　Ｄｕｐｌｅｘ）、ペア作成状態（Ｐａｉｒ　Ｃｒｅａｔｅ）、ペア中断状態（Ｐａｉｒ　Ｓｕｓｐｅｎｄ）にあるかどうかを示す情報を登録する。
【００３１】
記憶制御装置１０１は、コピーマネージャプログラム１１３を実行することによって、ペア作成（Ｐａｉｒ　Ｃｒｅａｔｅ）時にＭ－ＶＯＬ１３９の先頭アドレスから順次Ｒ－ＶＯＬにデータをコピーすることで、Ｍ－ＶＯＬに格納されたデータをＲ－ＶＯＬにバックアップする。さらに記憶制御装置１０１は、差分情報１２３を参照して、差異が有る部分のデータをＭ－ＶＯＬからＲ－ＶＯＬにコピーしたり、逆に、差分情報１２３を参照して、差異があるデータをＲ－ＶＯＬからＭ－ＶＯＬへコピーするようにリモートＩ／Ｏs処理プログラム１１２に指示を出す。ここで差分情報１２３とは、Ｍ－ＶＯＬとＲ－ＶＯＬのコピーデータの不一致を防ぐために、記憶装置１０９の記憶領域に対応して設けられたビットマップであり、リモートコピーした後に記憶装置１０９の記憶領域のデータが書き換えられた場合、その更新された領域に対応してビットフラグを立てて管理する。
また、ペア管理プログラム１１４は、ホスト１からのリストア要求に基づいて、指定された記憶装置１０９のデータをリストアする。尚、リストア処理の詳細は後述する。
【００３２】
以上、マスタサイト側の記憶装置システム３の構成について述べた。
一方、リモートサイト側の記憶装置システム４も、基本的には同様のプログラム及び情報を所有する。主な相違点を述べると、記憶装置２２は、マスタサイトの記憶装置２１のコピーデータを記憶するが、記憶装置２３はマスタサイト側に発生したログデータを記憶する。もちろんマスタサイト側に発生したログデータは、マスタ側の記憶装置に記憶するようにしても良いが、図２の例では、そのログデータを通信路６を介してリモートの記憶装置システム４に送信して記憶装置２３内の論理記憶装置１４０に格納するように構成される。尚、記憶装置２２，２３内の論理的記憶装置（ボリューム）１３９、１４０の構成は同様である。
【００３３】
不揮発メモリ１３７に格納される情報に関して言うと、コマンド処理プログラム１４２、リモートＩ／Ｏ処理プログラム１４３、コピーマネージャプログラム１４４、ペア管理プログラム１４５、ネットワーク処理プログラム１５９、記憶装置Ｉ／Ｏ処理プログラム１６２、システム構成管理プログラム１６３の各プログラム、及び記憶装置管理情報１５０、１５１、ボリューム構成情報１５２、ペア管理情報１５３、差分情報１６４、等々は、マスタサイト側の対応するプログラム或いは情報と同様である。
しかしながら、リモートサイトの記憶装置２３にログデータ及びその管理情報を格納し、管理情報に基いてログデータを復元処理すると言う特徴的な動作を行うために、不揮発メモリ１３７には、ログボリュームグループ管理プログラム１４６、使用量管理プログラム１４７、ログＩＤ管理プログラム１４８、Ｍａｒｋ　ＩＤ管理プログラム１４９、ログボリュームグループ構成情報１５４、ログボリュームグループ使用量管理情報１５５、ボリュームプール構成情報１５６、ログＩＤ管理情報１５７、Ｍａｒｋ　ＩＤ管理情報１５８が記憶される。これらのプログラム及び情報の機能、意味については、以後順次説明される。
【００３４】
次に、図３を参照してホスト１の構成について説明する。
ホスト１を例にして説明するが、ホスト２の構成も実質的にホスト１と同様である。
ホスト１は、所定のプログラムを実行するプロセッサ２０１、プロセッサ２０１が実行するＯＳやＡＰ及びＡＰが使用するデータを格納するために使用されるメモリ２０２、ＯＳやＡＰ、ＡＰが使用するデータが格納されるディスク装置２０７、通信路５を接続するＩ／Ｏアダプタ２０５、通信路１０を接続するネットワークアダプタ２０６、フロッピー（登録商標）ディスク等の可搬記憶メディアからのデータの読み出し等を制御するリムーバ

ブル記憶ドライブ装置２０９、液晶表示装置のような表示器２０３、キーボード或いはマウスのような入力器２０４、及びこれらの構成ユニットを接続し、ＯＳやＡＰのデータや制御データの転送に用いられるＬｏｃａｌ　Ｉ／Ｏネットワーク２０８とを有して構成される。

【００３５】

リムーバブル記憶ドライブ装置２０９で使用される可搬記憶媒体としては、ＣＤ－ＲＯＭ、ＣＤ－Ｒ、ＣＤ－ＲＷ、ＤＶＤやＭＯ等の光ディスク、光磁気ディスクや、ハードディスクやフロッピーディスク等の磁気ディスク等がある。尚、以下に説明される各プログラムは、可搬記憶媒体からリムーバブル記憶ドライブ装置２０９を介して読み出されることで、あるいはネットワーク４又は５を経由することで、ホスト１のディスク装置２０７にインストールされる。尚、ホスト１は、冗長性確保のために、プロセッサ２０１等の構成ユニットを複数有していても良い。

【００３６】

次に、ホスト１で実行されるプログラムの例について説明する。

これらのプログラム２１０は、ホスト１のディスク装置２０７又はメモリ２０２に格納され、プロセッサ２０１で実行される。

ホスト１は、ＯＳ２３０の下で動作するＡＰ２３３、データベースマネジメントソフトウエア（以下「ＤＢＭＳ」）２３２を有する。ＤＢＭＳ２３２は、ＯＳ２３０、ファイルシステム（ＦＳ）２３１、ボリュームマネージャ（ＶＭ）２３５等を介して記憶装置システム３にアクセスする。また、ＤＢＭＳ２３２は、ユーザが使用する他のＡＰ２３３との間で、トランザクション処理等のＩ／Ｏ処理のやり取りを行う。情報処理の性能を上げる目的で、ホスト１のメモリ２０２を用いてプロセッサ２０１が情報処理を実行する。

【００３７】

また、ホスト１が入出力を行う複数の論理記憶装置間で相関があるものをログボリュームグループ管理情報２４８として保持する。また、ペア管理情報２４９は、ホスト１が入出力を行う論理記憶装置であるＭ－ＶＯＬと、それとペアになるＲ－ＶＯＬの管理情報を保持する。ペア操作プログラム２４７は、記憶装置システム３に対してペア操作（Ｐａｉｒ　Ｃｒｅａｔｅ（ペアの作成）、Ｐａｉｒ　Ｓｕｓｐｅｎｄ（ペアの中断）、Ｐａｉｒ　Ｒｅｓｙｎｃ（ペアの再開）、Ｐａｉｒ　Ｄｅｌｅｔｅ（ペアの削除））を行う。ログ操作プログラム２４６は、ログの操作、例えばログ取得開始、ログ取得終了等を行う。ペア管理プログラム２４５は、記憶装置システム３のペアの状態を監視し、障害が無いか監視する。容量管理プログラム２４４は、記憶装置システム３の容量を情報として保持しログボリュームグループに含まれる論理記憶装置の全容量を算出する。マーク処理プログラム２４３は、ＡＰ２３３やＤＢＭＳ２３２がファイルをセーブした場合や閉じた場合に、コミットした処理の後に呼び出され、マークデータを作成して記憶装置システム３に書き込む処理を行う。マークＩＤ管理プログラム２４２は、マーク処理プログラム２４３がマークデータをライトする場合にタイマとマークＩＤ情報２４１を用いてマークデータを作成する。マークＩＤ情報２４１はログＩＤに対する管理情報であり、マークＩＤとログの時刻情報から成る。

【００３８】

次に、図４、図５を参照して、記憶装置管理情報１１５、１５０、及び記憶装置管理情報１１６、１５１について説明する。尚、符号は記憶装置システム３側を参照して説明する。

図４において、記憶装置管理情報１１５は、ホスト１に提供される記憶装置２１に関するアドレスを登録するエントリ３０１と、記憶装置システム３で記憶装置２１を統一的に識別するための論理的なアドレスを登録するエントリ３０２を有するテーブルである。

エントリ３０１は、ホスト計算機に提供される記憶装置の識別子を登録するエントリ３０３、及びその内部アドレスを登録するエントリ３０４を有する。また、エントリ３０２は、記憶装置システム３で記憶装置を統一的に識別する記憶装置２１内の論理記憶装置１０９の識別子３０５及びその内部アドレスを登録するエントリ３０６を有する。

【００３９】
図５において、記憶装置管理情報１１６は、記憶装置システム３内で記憶装置２１を統一的に識別する論理的はアドレスを登録するエントリ４０１と、ＲＡＩＤ　Ｇｒｏｕｐに関するアドレスを登録するエントリ４０２と、ＲＡＩＤ　Ｇｒｏｕｐを構成するディスクに関するアドレスを登録するエントリ４０３を有するテーブルである。
更にエントリ４０１は、記憶装置システム３内で論理記憶装置１０９を統一的に識別する論理記憶装置番号を登録するエントリ４０４と、記憶装置に対応する内部アドレスを登録するエントリ４０５を有する。エントリ４０２は、記憶装置システム３内で論理記憶装置１０９を統一的に識別する論理記憶装置番号によって識別される論理記憶装置が配置されているＲＡＩＤ　Ｇｒｏｕｐを記憶装置システム３内で統一的に識別するＲＡＩＤ　Ｇｒｏｕｐ番号を登録するエントリ４０６と、ＲＡＩＤ　Ｇｒｏｕｐを１つの記憶領域として扱った仮想空間の対応するアドレスを登録するエントリ４０７を有する。また、エントリ４０３は、ＲＡＩＤ　Ｇｒｏｕｐを形成するディスクを記憶装置システム３内で統一的に識別するディスク番号を登録するエントリ４０７と、ディスク内部アドレスが登録されるエントリ４０８を有する。
【００４０】
次に、図６を参照して、図２に示すＶｏｌｕｍｅ構成情報１１７、１５２について説明する。尚、Ｖｏｌｕｍｅ構成情報１５２も同様であるので、符号は記憶システム１側の符号を参照して記載する。
Ｖｏｌｕｅｍ構成情報１１７は記憶装置システム３におけるＶｏｌｕｍｅの構成に関する情報を登録するテーブルである。エントリ５０１には、記憶装置システム３内の論理記憶装置を統一的に扱う論理記憶装置番号が登録される。エントリ５０２には、ホストタイプが登録される。すなわち論理記憶装置に対して入出力を行うホストのＯＳが認識できる記憶装置のいずれに擬似化されている（エミュレートされる）かを示す情報、例えば、オープン系システムのＯＳが認識できる記憶装置であることを示す「ＯＰＥＮ」や、メインフレーム系のＯＳが認識できる記憶装置であることを示す「３９９０」等の情報が登録される。
【００４１】
エントリ５０３には、ホスト１が入出力を行えるように入出力ポートに関連付けられているかどうかを示すパス定義情報が登録される。例えばＩ／ＯネットワークがＦＣだったら、論理記憶装置とＦＣのＰｏｒｔとの関連付けに関する情報が登録される。
エントリ５０４は、論理記憶装置の状態を登録するものであり、例えば論理記憶装置に障害が無く入出力が行えるノーマル状態（ＮＯＲＭＡＬ）を示す情報や、障害などの理由で入出力が行えない状態（ＢＬＯＣＫＥＤ）を示す情報が登録される。更に障害情報としては、論理記憶装置が何らかの障害になったかどうかを示す情報が登録される。ここで、障害とは、主に論理記憶装置を構成する物理記憶装置の物理的障害や、管理者が意識的に記憶装置システムを閉塞状態にした場合等の論理的障害が含まれる。
【００４２】
エントリ５０５はリザーブ情報を登録するものであり、例えば論理記憶装置が、Ｒ－ＶＯＬやログデータを格納するために予約されている状態にあるかを示す情報が登録される。リザーブ情報が登録されている論理記憶装置は、その他の用途、例えば新たに業務用論理記憶装置として割り当てるなどが出来ない。
論理記憶装置１０９がペアを形成している場合には、ペア情報を登録するエントリ５０６にペア番号が登録され、記憶装置内でのログデータを取得する場合にはログボリュームグループ番号を登録するエントリ５０７にログボリュームグループ番号が登録される。エントリ５０６に、有効な情報例えばログボリューム番号が在れば、論理記憶装置がログ取得の対象、すなわちジャーナルモードの対象であるかどうかを示す情報が登録される。
またエントリ５０８には論理記憶装置１０９の容量が登録される。
【００４３】
次に、図７を参照して、ペア管理情報１１８、１５３について説明する。尚、参照符号

はペア管理情報１１８側について示す。
ペア管理情報１１８は、ペアを構成する正副の論理記憶装置の識別子を登録するテーブルである。すなわち、エントリ６０１にはペア番号が登録される。エントリ６０２には、ペアを形成するマスタ側の記憶装置システム３内の統一的な論理記憶装置（Ｍ－ＶＯＬ）番号が登録される。エントリ６０３には、Ｍ－ＶＯＬとペアを形成するリモート側の記憶装置システム４に在る論理記憶装置（Ｒ－ＶＯＬ）を識別する情報が登録される。例えばエントリ６０３には、通信路６で接続される入出力用ポートの識別子とこのポートからアクセス可能な記憶装置システム４内で統一的な記憶装置の番号が登録される。エントリ６０４には、ペアの状態、例えばペア状態、ペア作成中、ペア中断中、ペア再形成中を示す情報が登録される。

【００４４】

次に、図８以降の図面を参照して、リモート側の記憶装置システム４側に特有なテーブルの構成について説明する。
図８は、ログボリュームグループ構成情報１５４のテーブルの構成を示す図である。ログボリュームグループ構成情報１５４は、ログボリュームグループごとに用意される。カラム７０１にはログボリュームグループ毎を識別する情報が格納される。カラム７０２には、ログボリュームグループにグルーピングされた論理記憶装置の総数が格納される。カラム７０３には、ログボリュームグループにグルーピングされた論理記憶装置の容量の総和が格納される。

【００４５】

カラム７０４には、ログボリュームグループにグルーピングされた論理記憶装置の識別子７０５と、ＨＯＳＴ　ＴＹＰＥ７０６と、状態７０７と、容量７０８がログボリュームグループにグルーピングされた論理記憶装置の数分だけ登録される。カラム７０９には、ログボリュームグループで使用されるＬｏｇ用Ｖｏｌｕｍｅの論理記憶装置の識別子の数が登録され、カラム７１０にはログボリュームグループで使用するＬｏｇ用Ｖｏｌｕｍｅの論理記憶装置の容量の総和が登録される。
カラム７１０には、ログボリュームグループで使用するＬｏｇ用Ｖｏｌｕｍｅの論理記憶装置の識別子７１１と、ＨＯＳＴ　ＴＹＰＥ７１２と、状態７１３と、容量７１４がログボリュームグループで使用するＬｏｇ用Ｖｏｌｕｍｅの論理記憶装置の数分登録される。

【００４６】

次に、図９を参照してＶｏｌｕｍｅ　Ｐｏｏｌ構成情報１５６のテーブルの構成について説明する。
Ｖｏｌｕｍｅ　Ｐｏｏｌ構成情報１５６は、ログボリュームグループで使用するログデータを格納する記憶装置の容量が足りなくなった場合に容量を拡張しやすいように、予めパスが定義されていないものやリザーブ状態の論理記憶装置を一括管理するためのテーブルである。
エントリ８０１には、記憶装置システム４の論理記憶装置番号が登録される。エントリ８０２には、エントリ８０１の番号で識別される論理記憶装置のＨＯＳＴ　ＴＹＰＥが登録される。エントリ８０３には、同じくその論理記憶装置の状態が登録され、エントリ８０４には、その論理記憶装置のリザーブ情報が登録される。エントリ８０５にはその論理記憶装置の容量が登録される。

【００４７】

図１０は、ログボリュームグループ使用量管理情報１５５のテーブルの構成を示す。
ログボリュームグループ使用量管理情報１５５は、ログボリュームグループごとに、ログデータを格納する記憶装置の使用量を監視するために用意される。エントリ９０１にはログボリュームグループ番号が登録され、エントリ９０２にはログボリュームグループ内の全空き容量が登録される。

【００４８】

エントリ９０３には、ログボリュームグループで使用するログ用ボリュームの論理記憶装置の番号が格納される。エントリ９０４にはエントリ９０１で識別される論理記憶装置

の容量情報が格納される。エントリ９０５にはライトＩ／Ｏで使用された領域の容量が登録される。この領域の容量は、使用監視情報９０８のテーブルを参照して把握できる。すなわち使用監視情報９０８は、ログ用ボリュームの論理記憶装置１３９をビットマップで管理しており、ライト更新があった領域にビットを立てて管理する。この使用監視情報９０８のビットマップは記憶制御装置１３１の不揮発メモリ１３７内に格納されている。
エントリ９０６にはライト更新がかかっていない残り容量情報が格納される。エントリ９０７には、ログボリュームグループの全記憶容量に対する空き容量の割合を示す情報が格納される。

【００４９】

図１１は、ＬＯＧ　ＩＤ管理情報１５７のテーブルの構成を示す。
ＬＯＧ　ＩＤ管理情報１５７は、記憶装置システム３から送信されるライトＩ／Ｏによる更新データを記憶装置システム４で受信して、Ｉ／Ｏレベルのログを作成する時に、ログデータに付加される識別情報（ＩＤ）を管理するテーブルである。すなわち、ＬＯＧ　ＩＤ管理情報１５７のテーブルは、ログデータに付加されたログＩＤカウンタを管理するために、ログボリュームグループ毎に最も古いログデータのＩＤ値を格納するエントリ１００１と、そのデータが作成された時間情報を格納するエントリ１００２と、そのデータが格納されたログボリュームのアドレス情報を格納するエントリ１００３と、最も新しいログデータのＩＤ値を格納するエントリ１０１１と、そのデータが作成された時間情報を格納するエントリ１０１２と、そのデータが格納されたログボリュームのアドレス情報格納するエントリ１０１３から構成される。

【００５０】

コマンド処理プログラム１４２は、処理すべきコマンドがログデータの記憶される先の論理記憶装置１４０に対するライトＩ／Ｏ又はマークＩ／Ｏである場合に、ロググループ毎に用意されるログＩＤカウンタ１６０及びタイマ１６１を参照してカウンタ１６０の値（ログデータＩＤ）及びタイマ１６１の値（時間）を付加してログＩＤ管理情報を作成する。

【００５１】

次に、図１３を参照して、ログデータのフォーマットについて説明する。
ログデータは、記憶装置システム４が記憶装置システム３からのライトＩ／Ｏもしくはマークライト／Ｏ処理要求を処理する毎にキャッシュメモリ１３６上に作成され、その後記憶装置１４０に格納される。ログデータは、その先頭にマークフラグ１２０１が付加される。マークフラグ１２０１は、ホスト１と記憶装置システム４でシステムの状態を一意に識別するＭＡＲＫ情報であるかを識別する識別子である。マークフラグ１２０１の後には、ログＩＤ（即ちログＩＤカウンタの値）１２０２と、タイマ値１２０３と、データ長１２０４のそれぞれのエントリと、更にライトデータ又はマークデータが格納されるエントリ１２０５を備えて構成される。

【００５２】

マークデータの場合には、ログデータを格納するログボリュームのアドレス１２１０を格納するエントリと、ホスト１がマークＩＤ１２１１を格納するエントリと、マークを実行した時間１２１２を格納するエントリと、マークを要求したアプリケーションの識別情報１２１３が格納されるエントリより構成される。

【００５３】

一方、ライトデータの場合には、ライトデータを格納するログボリュームのアドレス１２２０を格納するエントリと、ライトデータ１２２１を格納するエントリより構成される。
尚、ログデータの作成については図２０を参照して後述する。

【００５４】

次に、図１２を参照して、ＭＡＲＫ　ＩＤ管理情報１５８のテーブルの構成について説明する。
ＭＡＲＫ　ＩＤ管理情報１５８は、ＭＡＲＫ　ＩＤ情報を管理するテーブルであり、ログ

ボリュームグループ毎に最も古いマークシーケンス番号の値を格納するエントリ１１０１と、そのデータを作成した時間情報を格納するエントリ１１０２と、そのデータが格納されたログボリュームのアドレス情報格納するエントリ１１０３と、最も新しいマークデータのＩＤ値を格納するエントリ１１０４と、そのデータを作成した時間情報を格納するエントリ１１０５と、そのデータが格納されたログボリュームのアドレス情報格納するエントリ１１０６を備えて構成される。

【００５５】

コマンド処理プログラム１１１が、ホスト１が送信されるマークコマンドを受領すると、このマークコマンドを使用してマーク情報を記憶装置システム３に送信してきたものを受領する。マークコマンドであれば、図２０に示すようにマーク情報をログデータ化するが、その際に、データ内のマークＩＤ１２１１を読み出してテーブルに格納して管理する。

【００５６】

次に、フローチャートを参照しながらそれぞれの処理動作について説明する。
まず、全体の処理動作の概要について説明しておく。
本実施形態による計算機システムでは、記憶装置システム４において、記憶装置システム３の正論理記憶装置（Ｍ－ＶＯＬ）のある時点のデータのバックアップデータ（以下「スナップショットデータ」）を有する副論理記憶装置（Ｒ－ＶＯＬ）を作成し、保持する。スナップショットデータが作成された時点以降に、ホスト１から受領するライトＩ／Ｏ処理要求を実行する度に、記憶装置システム３で発生した更新データを記憶装置システム３から記憶装置システム４に送信する。記憶装置システム４では、ライトＩ／Ｏ処理後のデータ（即ちライト更新データ）をログデータとして記憶装置２３に記録する。

【００５７】

さらに、ホスト１は、自らが作成する任意の情報であるマークポイント情報（以下「ＭＰ情報」）、即ちマークを付加する時点の情報を記憶装置システム３に対して通知する。具体的には、ホスト１は、任意の時点、例えば記憶装置システム３と記憶装置システム４との間でのデータを一致させる処理（シンク処理）を行う時に、ＭＰ情報を記憶装置システム４のログデータに書込む。従って、ＭＰ情報は、ホスト１と記憶装置システム４の両方で管理されることになる。これによって、ホスト１が指示するＭＰ情報及び記憶装置システム４内のログデータに格納されたＭＰ情報を利用して、記憶装置システム４は、ホスト１が意図した時（ＭＰ情報作成時）から記憶装置システム４が保持していたデータを高速にリストアすることができる。

【００５８】

このような処理を実行するために、ホスト１は、あらかじめ、ログデータを取得する準備指示（ログ取得開始準備指示）、及びログ取得開始指示を記憶装置システム４に送信する。これにより、記憶装置システム４は、ログデータの取得を開始し、ログモードとなる。その後、計算機システムは、上述したＭＰ情報の遣り取りを行う。

【００５９】

まず図１４を参照して、ペア作成の処理動作について説明する。
記憶装置システム３は、ホスト１からペアの作成を指示するコマンドを受領する（１３０１）。するとコマンド処理プログラム１１１は、ペア管理プログラム１１４を呼び出し、ペア作成要求コマンドの内容を渡す。ペア管理プログラム１１４は、コマンド情報を参照して、ペア作成指示対象のデバイス（記憶装置）のチェックを行う（１３０２）。
ペア管理用の差分管理ビットマップ１２３のビットを全て「１」にして、コピーを行うためにコピーの進捗を管理するためのポインタを用意し、ペア管理差分ビットマップ９０８の先頭を設定する（１３０３，１３０４）。

【００６０】

ペア対象の記憶装置として、リモートの記憶装置システム４が指定されている場合には、リモートの記憶装置の検出を行い、通信路の確立を実行する（１３０５）。リモートの記憶装置の検出が出来ない場合（１３０６）には、リトライ処理を実行する（１３０７）

。そして最終的に、リモートの記憶装置の検出が出来なかった場合には、障害が起きたものとしてホスト１に応答する（１３０９）。尚、システムの設計上予め記憶装置システム３と記憶装置システム４の間の通信路が確立している場合には、この一連処理は必要なくなる。

【００６１】

リモートの記憶装置が検出される（１３０６）と、リモートの記憶装置との通信路の確立処理が行われる（１３１０）。その後、ペア管理情報の設定と通信路の確立が出来たら（１３１１）、記憶装置システム４に対してペア作成コマンドを送信する（１３１６）。そして記憶装置システム４からペア作成コマンドに対する応答を記憶装置システム３が受領したら（１３１７，１３１８）、コピーマネージャプログラム１１３はコピー範囲を設定して（１３１９）、リモート側の記憶装置システム４に送信するデータを作成する。そしてリモートＩ／Ｏ処理プログラム１１２を実行してリモート側にそのデータを送信し、コピー進捗ポインタを進める（１３２０）。記憶装置システム４からデータ送信に対する完了報告が帰ってきたら、送信部分のビットマップ１２３を「０」クリアする（１３２１）。

そして、差分ビットマップ１２３が全て「０」かを判定して、もし全てがクリアされていれば（１３２２）、Ｐａｉｒ管理テーブル１１８のペア状態６０４を“ＤＵＰＬＥＸ”に登録して（１３２３）、ホスト１にペア作成完了報告を返す（１３２４）。

【００６２】

尚、通信路の確立や記憶装置システム３と記憶装置システム４間でのコマンド、データのやり取りで、記憶装置システム４からの状態の確認要求、エラー報告等の不具合が起きた旨の情報が応答として返って来た場合や、コマンドに対する応答時間が長く、タイムオーバした場合などはホスト１にエラー報告を返す。

【００６３】

次に、図１５を参照して、ペア作成中のコマンド処理動作について説明する。

ペア作成中にも記憶装置システム３はホスト１からデータの入出力及びその他システムの状態等を監視するコマンド等を随時受領している。この際のＩ／Ｏ処理要求を処理するコマンド処理プログラム１１１の動作を説明する。

【００６４】

コマンドを受領したコマンド処理プログラム１１１は、ライトＩ／Ｏ処理要求かを判断する（１４０１，１４０２）。この判断の結果、ライトＩ／Ｏの場合には、ペア作成対象デバイス（記憶装置）か否かを確認する（１４０３）。もしペア作成対象のデバイスであれば、ライトＩ／Ｏ対象アドレスとペア差分管理ポインタのアドレスとの比較を行い（１４０４）、ライトＩ／Ｏを送信済みかどうか確認する（１４０５）。もし送信済みであれば、ライトＩ／Ｏデータをキャッシュメモリ１０６の所定のエリアを確保して格納する（１４０６）。そしてペア管理プログラム１１４を呼び出して、このライトＩ／Ｏデータを格納したキャッシュメモリ１０６のそのエリアの情報を渡す。ペア管理プログラム１１４は、優先的に当該ライトデータに対するリモート側への送信処理を行うために、ペア情報を用いて、リモートへのＩ／Ｏ処理コマンドとデータを作成して、リモートＩ／Ｏプログラムを用いてリモートの第２の記憶装置システム４にデータを送信する（１４０７）。記憶装置システム４からライト完了報告応答を受領したら（１４０８）、ホストにライト処理終了報告を返す（１４０９）。

【００６５】

一方、リモートペア対象外の場合（１４０３Ｎ）や、ペア作成のためのデータ送信が未だ送信されていない場合（１４０５Ｎ）には、キャッシュメモリ１０６のエリアを確保してデータを格納し（１４１１）、その後ホスト１にライト処理終了報告を返す（１４０９）。

また、ライトＩ／Ｏか否かの判定（１４０２）においてライトＩ／Ｏ処理以外の場合には、所定の処理を行い、ホスト１に処理終了報告を返す。例えばリード処理の場合には、リードＩ／Ｏの対象アドレスに対応するキャッシュメモリ又は記憶装置からデータを読み出

して（１４２１）、ホスト１にリードデータを送信した後、完了報告を返す（１４２２）。

【００６６】

次に、図１６を参照して、マスタ側のログの取得開始処理について説明する。記憶装置システム３における記憶装置２１と、第２の記憶装置システムにおける記憶装置２２との間でペアを確立してＤＵＰＬＥＸ状態になったら、ホスト１は、任意の契機で、被災のために誤ったデータがリモートに送信されても、リモート側のペアに即座に反映されないように、ペア中断コマンドとログデータ取得開始コマンドを送信する。

【００６７】

記憶装置システム３は、ホスト１から送信されるペア中断（サスペンド）コマンド及びログデータ取得開始コマンドを受領すると（１５０１）、その受領したコマンドを解析して、ペア管理プログラム１１４を呼び出して、ログボリュームグループ番号、ペア状態、デバイス（記憶装置）を確認し（１５０２）、リモートＩ／Ｏプログラム１１２を実行してリモートの記憶装置システム４にサスペンドコマンド及びログデータ取得開始コマンドを送信する（１５１１）。

【００６８】

続いて記憶装置システム３は、記憶装置システム４から送信されてくるログ取得準備完了の応答の受信を待つ。そして準備完了の応答を受領したら（１５０３）、記憶装置システム４側のペア管理プログラム１４５は、指示されたペアに関する記憶装置番号を基にペア構成情報１１８のペア番号を割り出し、ログボリュームグループ内のこのペアの状態をサスペンドに設定する（１５０４）。その後、ログ取得コマンドに対する完了報告をホスト１に返す（１５０５）。

一方、ログ取得準備完了の応答ではなく（１５０３）、不具合な応答を受領した場合（１５２１）には、ホスト１にエラー報告をして処理を終わる（１５２２）。

【００６９】

次に、図１７を参照して、リモート側のＬＯＧ取得開始処理について説明する。リモート側の記憶装置システム４は、記憶装置システム３からペアサスペンドコマンド及びログデータ取得開始コマンドを受領すると（１６０１）、受領したコマンドを解析して、ペア管理プログラム１４５を呼び出してログボリュームグループ番号、ペア状態およびデバイスを確認する（１６０２）。

続いて、ボリューム構成情報テーブル１５２を参照してペア確認を行い、関連するペア管理情報テーブル１５３をサスペンドにして（１６０３）、ログボリューム用記憶装置（リザーブ）が用意されているかを確認する。この時、ログボリューム用リザーブが用意されていなければ、論理記憶装置１４０を割り当てる。その後ログボリュームグループ用のログシーケンスカウンタを初期化し（１６０４）、ログボリュームグループ情報フラグをＯＮにして、完了報告を記憶装置システム３に応答する（１６０５）。

【００７０】

次に、図１８を参照して、ＭＡＲＫコマンドの処理動作について説明する。ホスト１と記憶装置システム４の両方で同一のマーク情報を保持するために、ホスト１から発行されたマークコマンドは、記憶装置システム３を通して記憶装置システム４に送信される。

【００７１】

図１８Ａに示すように、ホスト１側では、ホスト１のＡＰやＤＢＭＳ（データベース管理システム）は、ファイルのセーブ時、ファイルのクローズ時、ＡＰのセーブ時、ＡＰのプロセスコミット時などの後にマーク処理プログラム２４３を呼び出してマークコマンドを発行する（１７０１）。マークコマンドを発行する際にマーク処理プログラム２４３はマークカウンタとタイマを参照し、マークＩＤとタイマ値を挿入して、マークデータを作成する。そして記憶装置システム３からＲｅａｄｙ応答を受領すると（１７０２）、記憶装置システム３へマークデータを送信する（１７０３）。その後一連の処理が終了して、記憶装置システム３から完了報告を受領すると、マークコマンドの処理を終える（１７０

４）。

【００７２】

次に、図１８Ｂを参照して記憶装置システム３におけるＭＡＲＫコマンドの処理について説明する。

記憶装置システム３のコマンド処理プログラム１１１は、ホスト１から送信されるマークコマンドを受領して（１７０５）、そのコマンドを解析する（１７０６）。解析の結果、もしそのコマンドがマークコマンドならば、ホスト１へＲｅａｄｙ応答を返す（１７０７）。

【００７３】

記憶装置システム３は、ホスト１からコマンドに続いて送信されるマークデータ又はライトデータを受領する（１７０８）。そして受領したコマンドに応じた処理を行う（１７０９）。もし、マークコマンドの場合には、リモートの記憶装置システム４へマークコマンドを送信する（１７１０）。そして記憶装置システム４からＲｅａｄｙ応答を受領したら（１７１１）、記憶装置システム４へマークデータを送信する（１７１２）。マークデータを送信した後、記憶装置システム４ではマークデータの処理が行われる。そして、処理が終わると、記憶装置システム３へ完了報告を送信する。記憶装置システム３は、リモートの記憶装置システム４から完了報告を受領すると（１７１３）、記憶装置システム３は、ホスト１にマークコマンド処理完了報告を返す（１７１４）。

【００７４】

ところで、先のステップ１７０６の判定において、受領したコマンドがリードコマンドである場合には（１７１５）、まずキャッシュのヒット/ミスヒット、即ちキャッシュメモリ１０６に目的とするアドレスの有無を判定する（１７１６）。そして、キャッシュメモリ又は記憶装置２１から目的とするデータを読み出して（１７１８）、ホスト１へ送信する（１７１８）。一方、リードコマンドで無ければ所定の処理を行う（１７１９）。

また、上記ステップ１７０９の判定で、ライトコマンドならば、キャッシュのヒット/ミスヒットを判定し（１７２０）、ライトデータを対応する記憶装置１０９及びキャッシュメモリ１０６に書き込み、非同期処理のキューに登録する（１７２１）。そしてリモート側へ送信すべき対象のデータか否かを判定し（１７２２）、送信すべきデータならばステップ１７１０へ移り、以後の処理を行う。リモート送信データで無ければ、完了報告をホスト１へ送信して終了する。

【００７５】

次に、図１９を参照して、リモートサイトの記憶装置システム４におけるログ取得中のコマンド処理について説明する。

記憶装置システム４は、記憶装置システム３からコマンドを受領すると（１８０１）、マークコマンドかライトコマンドか、それともそれ以外かを判定する（１８０２）。マークコマンドを受領した場合には、マークコマンドに対するＲｅａｄｙ応答を記憶装置システム３に返す（１８０３）。

【００７６】

マスタ側の記憶装置システム３では、リモート側からＲｅａｄｙコマンドを受領すると、リモートにマークデータを送信する（１７１２）。リモート側の記憶装置システム４では、マークコマンドの受領に引き続いてマークデータを受領すると（１８０４）、ログ取得対象デバイス（論理記憶装置）か否かを判定する（１８０５）。この判定は、リモート側のボリューム構成情報１５２、ペア管理情報１５３、ログボリュームグループ構成情報１５４のテーブルを順に参照することにより行われる。この判定の結果、対象のデバイスがログデータ取得対象のデバイスであるならば、ログデータ作成処理を行う（１８０６）。尚、ログデータ作成処理については図２０を参照して後述する。ログデータの作成が完了すると、マスタ側の記憶装置システム３へ完了報告を返す（１８０７）。

【００７７】

ところで、ステップ１８０２の判定の結果、リードコマンドを受領した場合には、リードコマンドの処理（１８０８～１８１２）を行い、ライトコマンドのときはライトコマン

ドの処理（１８１３～１８１４）の行う。これらの処理は、前述した図１８に示す動作と同様であるので、説明を省略する。

【００７８】

次に、図２０を参照して、記憶装置システム４におけるログデータ作成処理について説明する。

ログデータ作成処理では、まずログＩＤ管理情報テーブル１５７を参照しＬＯＧ　Ｖｏｌｕｍｅのアドレスを設定し（１９０１）、ログＩＤカウンタ１６０及びタイマ１６１を参照してログＩＤとタイマ値を設定する（１９０２）。そしてログデータ格納用にキャッシュメモリ１３６内に領域を確保する（１９０３）。その後、図１３（Ａ）に示すようにマークデータ（マークＩＤ１２１１及びマーク時間１２１２等を含む）に、ログ情報（ログＩＤ１２０２及びログ時間１２０３等を含む）を付加したログデータをキャッシュメモリ１３６の領域に格納する（１９０４）。後で非同期処理にこのログデータ作成処理を追加する（１９０５）。即ち、非同期記憶装置Ｉ／Ｏ処理プログラムの実行によりキャッシュメモリ１３６の当該領域に格納されたログデータは、記憶装置２３内の指定された論理記憶装置１４０に格納される。その後記憶装置システム３に完了報告を返して終了する。

【００７９】

次に、図２１を参照して、ＬＯＧ　Ｖｏｌｕｍｅ容量オーバー処理について説明する。

ログボリュームグループ管理プログラム１４６は、ロググループ毎にログボリュームの使用量を監視する。ライトＩ／Ｏもしくはマーク Ｉ／Ｏによる更新が生じた場合には、図１０に示すログボリューム使用量管理情報１５５に反映される。更にこの際に反映したログボリューム使用量管理情報１５５を参照して（２００２）、ログボリュームの使用量が規定値以上の量を使用していないかを確認する（２００３）。尚、この規定値は、ホスト１又は管理装置８，９もしくは計算機システム管理装置１３を用いて設定されている。上記判定の結果、規定値以上の使用量であれば、通信路６を介して記憶装置システム３に対してログボリュームの容量不足の旨のＡｌｅｒｔを出す（２００４）。

【００８０】

このＡｌｅｒｔを受けた記憶装置システム３は、このＡｌｅｒｔをホスト１に返す。または、第２の記憶装置管理装置４へＡｌｅｒｔを出し、第２の管理装置９から計算機システム管理装置１３やホスト１にＡｌｅｒｔ情報が伝達されるようにする。この場合、システム管理者は、ログボリュームを反映するように記憶装置システム４に指示をする。ホスト１から指示を出す場合は、ホスト１のマークＩＤ管理情報を参照して、適当なマークＩＤを指示してログボリューム内のログデータをＲ－ＶＯＬに反映して、ボリューム容量不足を回避するか、容量追加を指示する。それを受領した記憶装置システム４は、Ｖｏｌｕｍｅ　Ｐｏｏｌ構成情報１５６に登録されている空きＶｏｌｕｍｅを割り当てることによりボリューム容量不足を回避する。

【００８１】

次に、図２２を参照して、Ｔａｋｅ　Ｏｖｅｒ処理について説明する。

この処理はホスト１、２と連携して行われる処理である。

まず、ホスト２のクラスタソフトウェアは、ホスト１のクラスタソフトウェアとの間で通信路１０を介してハートビートの監視を行う（２１０１）。具体的には、所定の時間間隔に、決まったあるデータをホスト１とホスト２のクラスタソフトウェアは通信しあい、ある所定の時間内での応答を期待する。

この監視において、ハートビートに対する応答が所定回数無い場合には、相手方ホストの障害かネットワーク通信路の障害が想定されるので、ハートビート切れと判断して（２１０２）、交替系の通信路を使用して更にハートビートを行う（２１０３）。尚、交替系の通信路は、通常その機能を満たすためには、正常系の通信路と物理的に分かれていたほうが望ましい。

【００８２】

交替系の通信路でも、応答がない場合には（２１０４）、システム管理者に障害が発生した可能性があるためＴａｋｅ　Ｏｖｅｒを行うための問い合わせメッセージを出す（２

１０４）。システム管理者は、ホスト１があるサイトで何らかの災害が発生して障害が起きた場合で、ホスト１がまだ計算機処理を行える状態であった場合は第１のクラスタソフトを用いてＴＡＫＥ　ＯＶＥＲ処理を実行する。具体的には、記憶装置システム３のリモートペアを全てサスペンドにするコマンドを記憶装置システム３に送信すると共に、第２のホストに計算機システム４の処理を交代する。この際に処理中のアプリケーションは中断され、リモート間のコピー処理は中断する（２１０６）。

【００８３】

ホスト２においても、記憶装置システム４に対して、サスペンド処理を行うコマンドを送信する。もしくはホスト１との間でのハートビートが、一定時間途絶えたら、ホスト２のシステム管理者はホスト１のシステム管理者に問い合わせ、ホスト１側が被災していないかを確認し、被災していることが判明した場合や、既にマスコミ媒体によって被災の情報が知れている場合には、ホスト２から記憶装置システム４にペアサスペンドコマンドを投げ、ペアを中断する。そしてペア情報はサスペンドに設定される（２１０７）。
しかし、被災中のためホストの情報処理のためのメモリの状態が異常になるなどで誤ったデータを書き込む処理を行う可能性がある。この場合も記憶装置システム３は、記憶装置システム４へデータを投げる。以後、ホスト１又は２は記憶装置システム４を用いてシステムのリカバリを行う。

【００８４】

次に、図２３を参照して、リカバリ処理について説明する。
このリカバリ処理とは、記憶装置システム４がホスト２からリカバリリストア指示を受領して行うデータ回復のための処理をいう。尚、以下の処理は、記憶制御装置１３１が、コピーマネージャプログラム１４４とログＩＤ管理プログラムを実行することで行われる（２２０１）。
ホスト１のリカバリ処理を行うホスト２にとって、論理的不整合等の障害が起きる可能性があるデータを、記憶装置システム４が記憶装置システム３から受け取っている可能性がある。そこで、ホスト２からリカバリ処理を行う場合に、記憶装置システム３の記憶装置Ｍ－ＶＯＬとペアを構成している記憶装置システム４のＲ－ＶＯＬ及びログ論理記憶装置１４０に格納されたデータを使用し、ホスト２から記憶装置システム４に対して、リストアすべき論理ボリュームの識別子とＲ－ＶＯＬとマークポイントを用いたリストア処理要求を送信する（２２０２）。

【００８５】

ホスト２は、リストア処理要求を受領する論理記憶装置を確認する（２２０３）。これは、記憶装置システム４に対して、マークＩＤ管理情報を取得するためのコマンドを送信し、マークＩＤ管理情報を取得することにより行う。取得されたマークＩＤ管理情報は、例えばホスト２の表示器、又は管理装置９の表示器に表示され、システム管理者によりリストア先の論理記憶装置が確認される（２２０４）。即ち、システム管理者は、表示されたマークＩＤ管理情報を参照し、回復ポイントを決める。そしてこれに対応するマークＩＤを選択し、回復コマンドの内容に含め送信するようにホスト２を用いて指示する。
ここで、ホスト２のシステム管理者は、障害発生直前のマークＩＤではなく、マークＩＤ情報のリストから、任意のマークＩＤを選択することができる。これにより、システムのユーザは、任意のマークＩＤを選択することで、選択されたマークＩＤが作成された時点に記憶装置システム４のＲ－ＶＯＬの状態を復元することが出来る。

【００８６】

ホスト２は、ステップ２２０１で選択したマークＩＤまでのログデータのリストア処理要求を記憶装置システム４に発行する（２２０５）。リカバリ処理となるリストア処理要求には、ログデータを反映する処理の対象となるＲ－ＶＯＬの識別子（例えばＷＷＮとＬＵＮ等）、Ｒ－ＶＯＬの属するログボリュームグループを指定する識別子、選択されたマークＩＤの情報等が含まれる。ホスト２より発行されたリストア処理要求を受領した記憶装置システム４の記憶制御装置１３１は、システム構成情報管理プログラム１６３を実行して、リストア処理要求に含まれるログボリュームの識別子とログボリュームグループ情

報を比較参照し、指定されたログボリュームがＲ－ＶＯＬに対する正しいログボリュームであるかを確認する。（２２０６）。

【００８７】

更に、記憶制御装置１３１は、リストア処理要求の内容から、Ｒ－ＶＯＬにログデータの反映処理を行うのか、或いは異なった未使用の記憶装置１０９にリストア処理を行うのかを確認する。尚、記憶装置１０９の障害により処理続行が出来ない場合はその旨をホスト２に通知し、処理を中止する。

記憶装置２１以外の他の記憶装置へデータをリストアする場合には、記憶制御装置１３１は、キャッシュメモリ１３６にデータ格納領域を確保する。その後、記憶制御装置１３１は、コピーマネージャプログラム１４４を実行して、Ｒ－ＶＯＬに対応するログボリュームの先頭から、順次ログデータをキャッシュメモリ１３６に確保された領域に読み出す（２２０７）。その際、読み出されたログデータにマーク情報が含まれるかどうかを確認する（２２０８）。即ちログデータのマークフラグがＯＮになっているかどうかを確認する。

【００８８】

読み出されたログデータが、マーク情報を含むログデータである場合、記憶制御装置１３１は、更に読み出されたログデータに含まれるマークＩＤがホスト２から指定されたマークＩＤかどうかを確認する（２２０９）。ログデータに含まれるマークＩＤがホスト２から指定されたマークＩＤでない場合、又はログデータにマークＩＤが格納されていない場合（マークフラグ（ＭＰ）がＯＮになっていない場合）には、記憶制御装置１３１は、読み出されたログデータに含まれるライトデータを、Ｒ－ＶＯＬ又はその他の記憶装置の対応するアドレスに書き込むように送信する。一方、マークＩＤに対応するログデータである場合には、ライトデータが存在しないので、データの書き込みは行われない（２２１２）。

その後、記憶制御装置１３１は、ステップ２２０７に戻り、次のログデータを読出し処理する。以下、記憶制御装置１３１は、２２０７～２２１１の処理を繰り返すことで、指示されたマークＩＤまでのログデータを指定された記憶装置のアドレスにリストアする。

【００８９】

上記ステップ２２０９で、マークＩＤが指定されたマークＩＤと一致した場合には、記憶制御装置１３１は、リストアすべきデータをすべてＲ－ＶＯＬや他の記憶装置１０９に書き込んだと判断して、リストア処理の終了をホスト２に通知する。

ホスト２は、記憶装置システム４から終了報告を受領したら、ホスト２が指定したマークＩＤ時点までのデータが回復されたと判断して、他の処理を継続する。記憶装置システム４は、ホスト２からリカバリ処理の完了応答を受領すると（２２１２）、処理を終了する。

【００９０】

以上、本発明の一実施例について説明したが、本発明は上記以外にも種々変形して実施し得る。

例えば上記実施例では、ログデータをリモートサイトの記憶装置システム４内に記憶するように構成している。しかし変形例では、リモートサイト側だけでなく、マスタサイトの記憶装置システム３内の記憶装置にもログデータを記憶するようにしてもよい。この場合には、図２に示した記憶装置システム４側に備えられたプログラムや情報は、マスタサイトの記憶装置システム３にも具備することは明らかある。

【００９１】

また、リストア先となる記憶装置は、必ずしもマスタサイトの記憶装置システム３内の記憶装置２１とは限らない。マスタサイトが被災状態にあるときには、リモートサイトの記憶装置システム４内にある記憶装置１３９内のボリューム１３９でもよい。或いはネットワークに接続された他の記憶装置システム内の記憶装置でもよい。

【００９２】

通常、リストア先の記憶装置にログデータをリストアした後は、マスタサイトの記憶装

置が保持していたデータが復元されたものとして扱われるので、その時点までのログＩＤ管理情報１５７、マークＩＤ管理情報１５８、及びログデータ等は、捨てられてよい。しかしながら、データの改ざん防止のため、或いは一定期間履歴として残しておくべき規則などがある場合には、上記管理情報及びログデータはそのまま保持し続けてもよい。

【図面の簡単な説明】

【００９３】

【図１】一実施形態による計算機システムの構成を示すブロック図。

【図２】記憶装置システム３及び４の構成を示すブロック図。

【図３】ホスト１の構成を示すブロック図。

【図４】記憶装置管理情報１１５、１５０のテーブルの構成を示す図。

【図５】記憶装置構成情報１１６、１５１のテーブルの構成を示す図。

【図６】Ｖｏｌｕｍｅ構成情報１１７、１５２のテーブルの構成を示す図。

【図７】ペア管理情報１１８、１５３のテーブルの構成を示す図。

【図８】ログボリュームグループ情報１１０のテーブルの構成を示す図。

【図９】Ｖｏｌｕｍｅ　Ｐｏｏｌ構成情報１５６のテーブルの構成を示す図。

【図１０】Ｌｏｇ　Ｖｏｌｕｍｅ使用量管理情報１５５のテーブルの構成を示す図。

【図１１】ＬＯＧ　ＩＤ管理情報１５７のテーブルの構成を示す図。

【図１２】ＭＡＲＫ　ＩＤ管理情報１５８のテーブルの構成を示す図。

【図１３】ログデータのフォーマットを示す図。

【図１４】ペア作成の処理動作の説明に供するフローチャートを示す図。

【図１５】ペア作成中のコマンド処理動作の説明に供するフローチャートを示す図。

【図１６】マスタ側のログ取得開始処理動作のフローチャートを示す図。

【図１７】リモート側のＬＯＧ取得開始処理動作のフローチャートを示す図。

【図１８Ａ】ＭＡＲＫコマンドの処理動作のフローチャート（ホスト側）を示す図。

【図１８Ｂ】ＭＡＲＫコマンドの処理動作のフローチャート（記憶装置システム３側）を示す図。

【図１９】リモートサイトにおけるログ取得中のコマンド処理のフローチャートを示す図。

【図２０】リモートサイトにおけるログデータの作成処理動作のフローチャートを示す図。

【図２１】ＬＯＧ　Ｖｏｌｕｍｅ容量オーバー処理動作のフローチャートを示す図。

【図２２】Ｔａｋｅ　Ｏｖｅｒ処理のフローチャートを示す図。

【図２３】リカバリ処理のフローチャートを示す図。

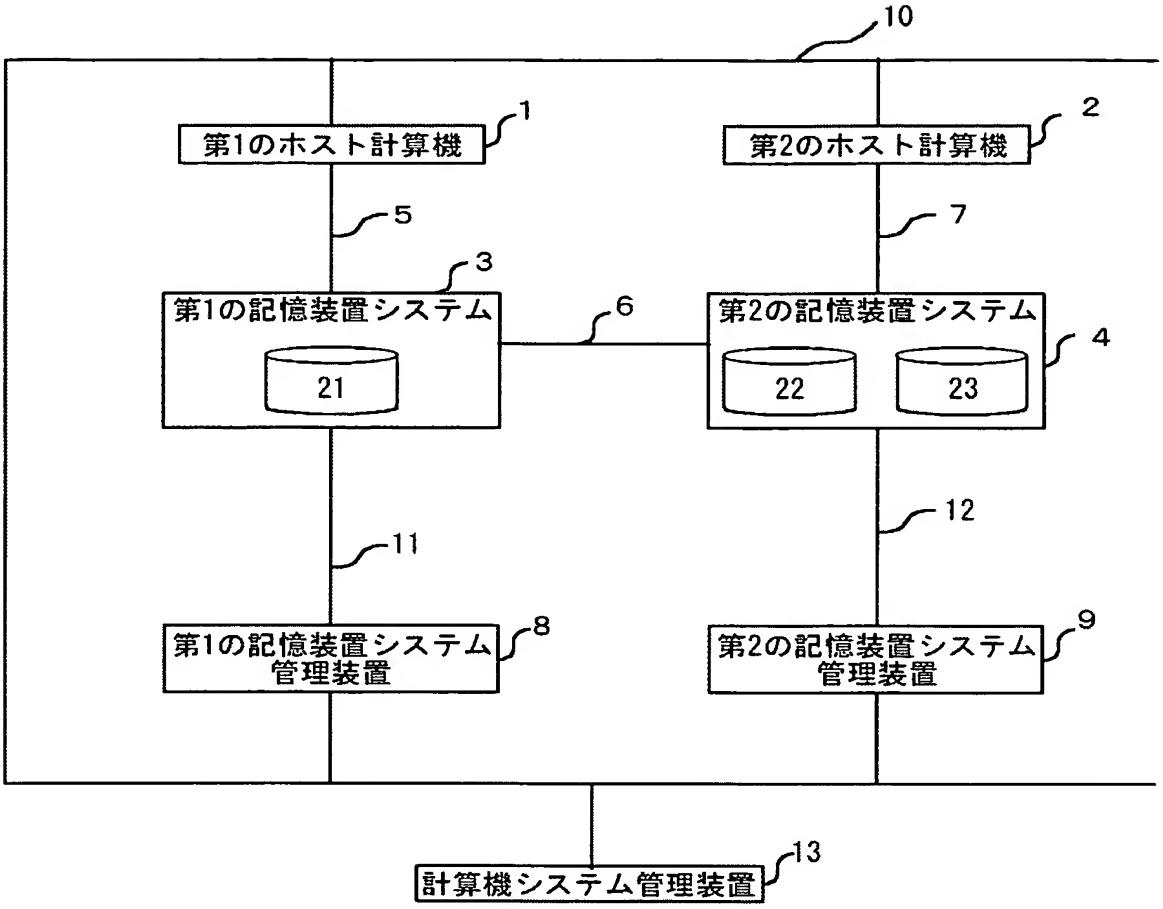
【符号の説明】

【００９４】

１：ホスト、　２：ホスト、　３：第１の記憶装置システム、　４：第２の記憶装置システム、　８，９：記憶装置システム管理装置、　１３：計算機システム管理装置、　５，６，７，１０、１１，１２：通信路、
２１，２２，２３：記憶装置、　１０９、１３９、１４０：記憶装置、

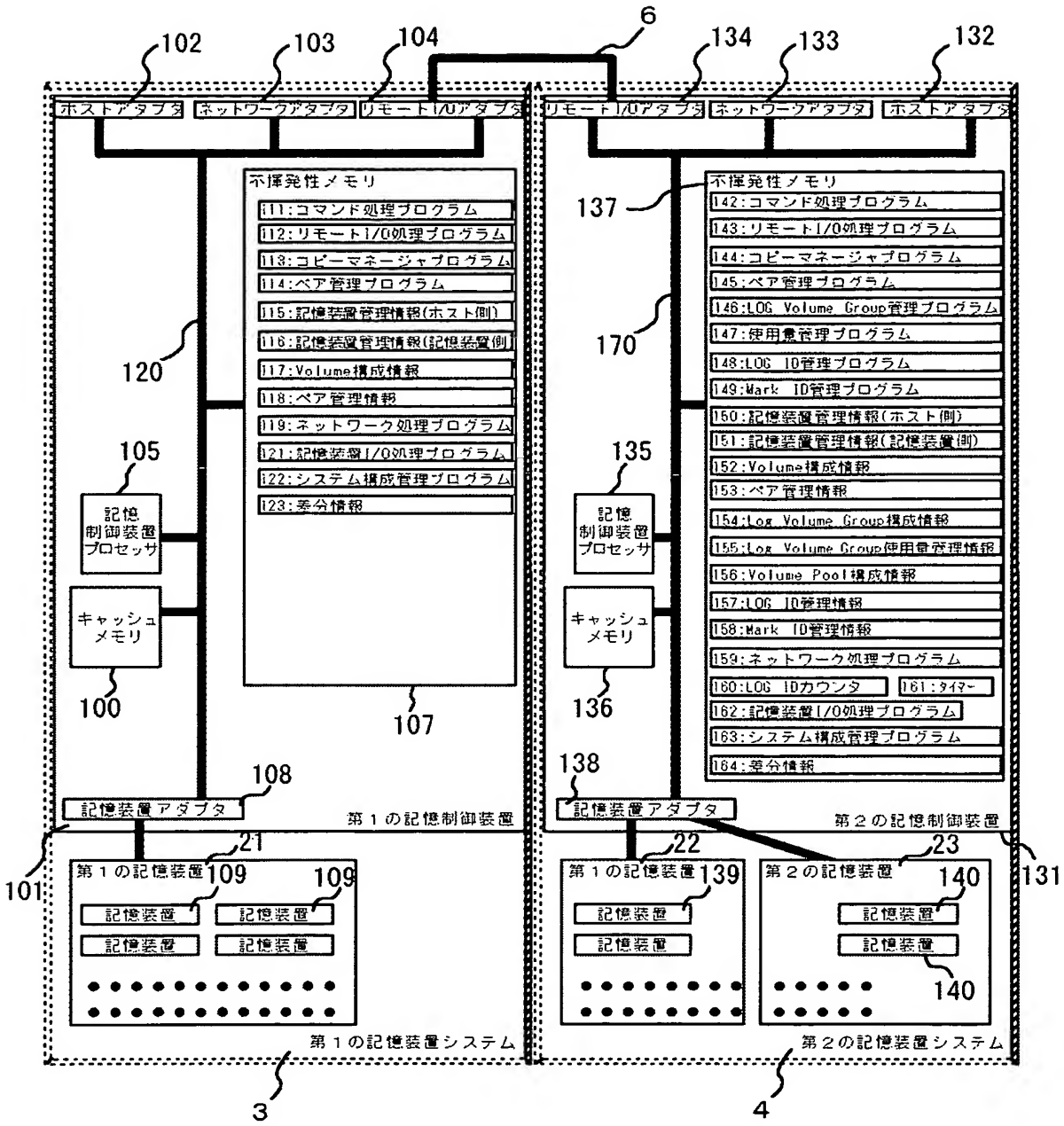
【書類名】図面
【図１】

図　1


10
第1のホスト計算機
1
第2のホスト計算機
2
5
7
3
第1の記憶装置システム
6
第2の記憶装置システム
21
22
23
4
11
12
第1の記憶装置システム
管理装置
8
第2の記憶装置システム
管理装置
9
計算機システム管理装置
13

【図２】

図　２


6
102
103
104
134
133
132
ホストアダプタ
ネットワークアダプタ
リモートI/Oアダプタ
リモートI/Oアダプタ
ネットワークアダプタ
ホストアダプタ
不揮発性メモリ
111:コマンド処理プログラム
112:リモートI/O処理プログラム
113:コピーマネージャプログラム
114:ペア管理プログラム
115:記憶装置管理情報(ホスト側)
116:記憶装置管理情報(記憶装置側)
117:Volume構成情報
118:ペア管理情報
119:ネットワーク処理プログラム
121:記憶装置I/O処理プログラム
122:システム構成管理プログラム
123:差分情報
120
105
記憶制御装置プロセッサ
キャッシュメモリ
100
107
108
記憶装置アダプタ
第１の記憶制御装置
101
21
第１の記憶装置
109
109
記憶装置
記憶装置
記憶装置
記憶装置
第１の記憶装置システム
3
137
不揮発性メモリ
142:コマンド処理プログラム
143:リモートI/O処理プログラム
144:コピーマネージャプログラム
145:ペア管理プログラム
146:LOG Volume Group管理プログラム
147:使用量管理プログラム
148:LOG ID管理プログラム
149:Mark ID管理プログラム
150:記憶装置管理情報(ホスト側)
151:記憶装置管理情報(記憶装置側)
152:Volume構成情報
153:ペア管理情報
154:Log Volume Group構成情報
155:Log Volume Group使用量管理情報
156:Volume Pool構成情報
157:LOG ID管理情報
158:Mark ID管理情報
159:ネットワーク処理プログラム
160:LOG IDカウンタ
161:タイマー
162:記憶装置I/O処理プログラム
163:システム構成管理プログラム
164:差分情報
170
135
記憶制御装置プロセッサ
キャッシュメモリ
136
138
記憶装置アダプタ
第２の記憶制御装置
131
22
第１の記憶装置
139
記憶装置
記憶装置
23
第２の記憶装置
140
記憶装置
記憶装置
140
第２の記憶装置システム
4

【図３】

図　３

ホスト計算機ブロック図

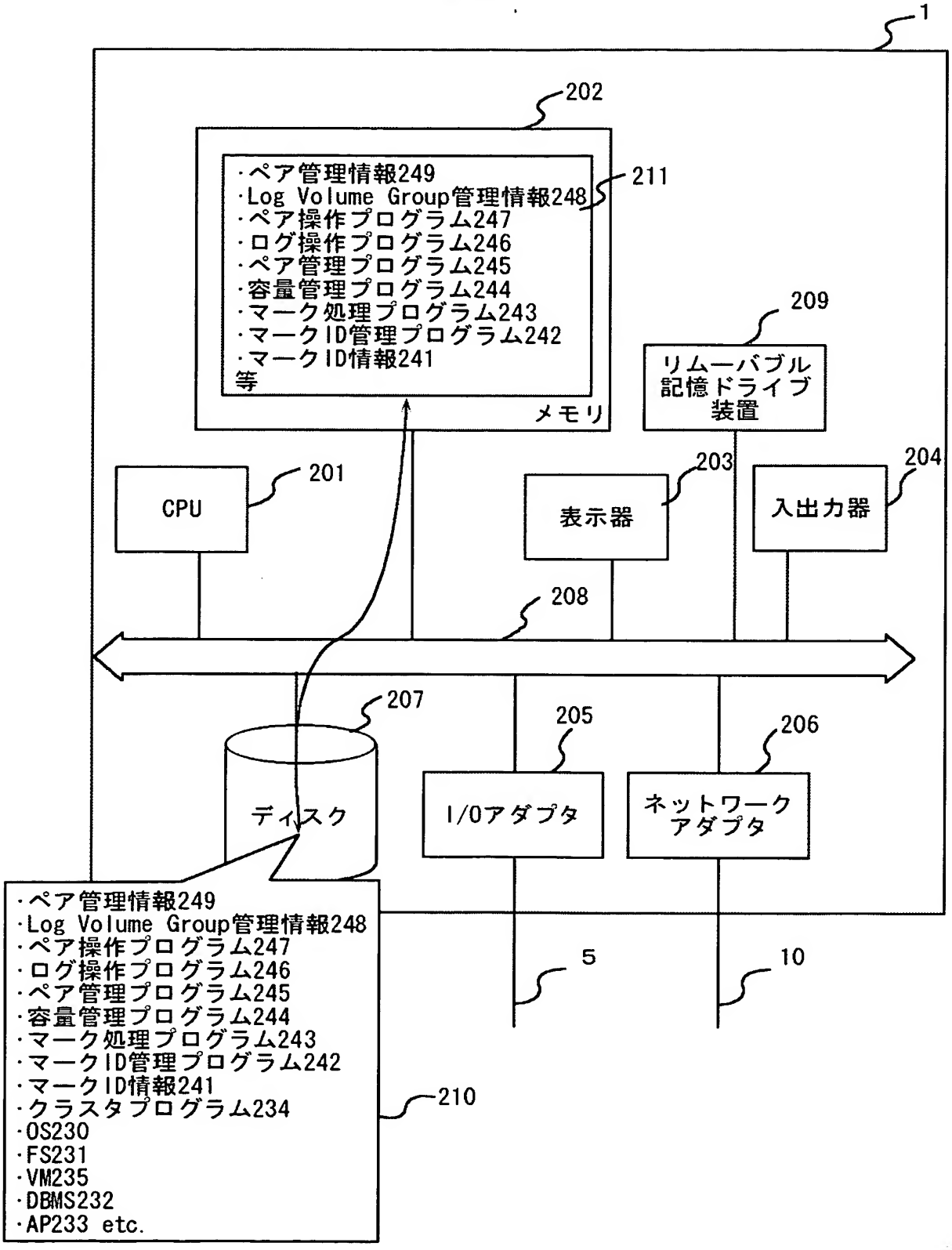


図　4

記憶装置管理情報 115

ホスト計算機の識別子：ＸＸＸＸＸＸＸＸ

ホスト計算機が認識する記憶装置のアドレス (301)		記憶装置システム内部での論理的なアドレス (302)	
記憶装置の識別子 (303)	記憶装置内アドレス (304)	記憶装置システム内論理記憶装置番号 (305)	記憶装置システム内論理記憶装置番号内アドレス (306)
aaaaaaaaaaaaaaaa	000000000000	00	000000000000
aaaaaaaaaaaaaaab	000000000000	0B	000000000000
aaaaaaaaaaaaaaab	100000000000	0C	000000000000
.....	.....	.....	-----------

【図５】

図　5

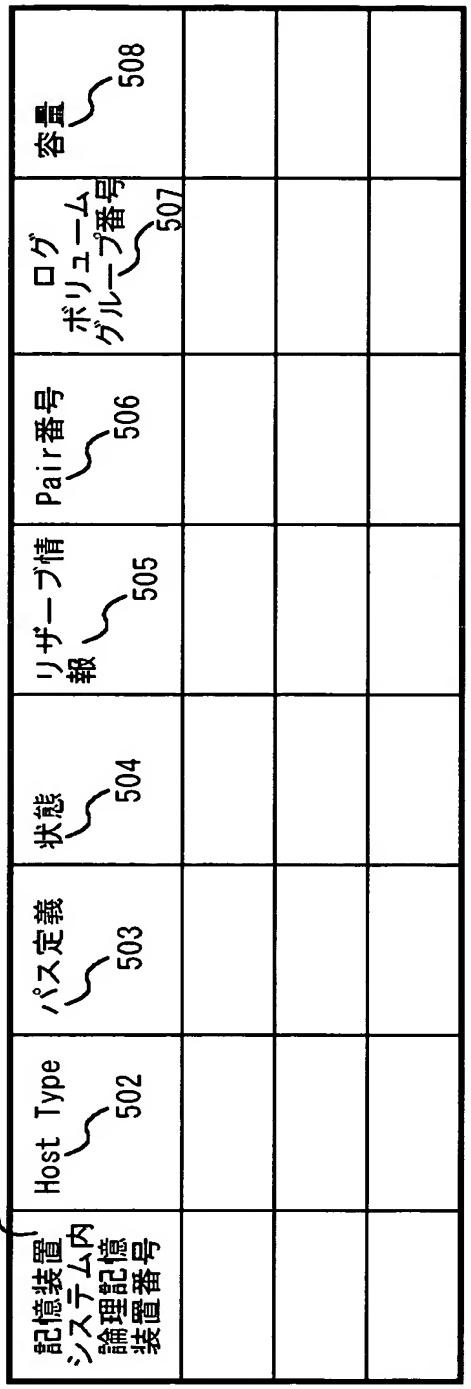
記憶装置管理情報　116

記憶装置システム内部での論理的なアドレス 401		RAID Groupに関するアドレス 402		ディスクに関するアドレス 403	
記憶装置システム内論理記憶装置番号 404	記憶装置システム内論理記憶装置番号内アドレス 405	RAID Group番号 406	RAID Group内仮想アドレス 407	ディスク番号 408	ディスクアドレス 409
00	000000000000	00	000000000000	00	000000000000
00	000002000000	00	000002000000	01	000002000000
.....	.....	.....	.....	.....	.....
....	....	....	....	....	....

【図６】

図　6

Volume構成情報　117,152



【図７】

図　７

Pair管理情報118.153

Pair番号 601	記憶装置システム3内論理記憶装置番号（正）602	記憶装置システム4内論理記憶装置番号（副）603	状態 604

【図８】

図　８

Log Volume Group構成情報　154

フラグ				701
Log Volume Group番号:xxx				702
論理記憶装置数:yy				703
Log Volume Group内論理記憶装置容量総和:zzz				704
記憶装置システム内論理記憶装置番号	Host Type 706	状態 707	容量 708	705
...	...	...	...	
Log Volume Groupで使用するLog用Volumeの論理記憶装置識別子数				709
Log Volume Groupで使用するLog用Volumeの論理記憶装置の容量総和				710
記憶装置システム内論理記憶装置番号	Host Type 712	状態 713	容量 714	711
...	...	...	...	

【図９】

図　９

Volume Pool構成情報

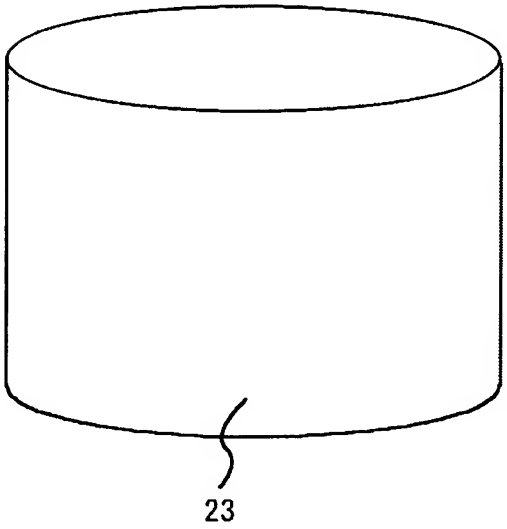
記憶装置システム内論理記憶装置番号 (801)	Host Type (802)	状態 (803)	リザーブ情報 (804)	容量 (805)

【図１０】

図　１０

Log Volume Group使用量管理情報　155

Log Volume Group番号 (901)				
Log Volume Group内全空き容量 (902)				
記憶装置システム内論理記憶装置番号 (903)	容量 (904)	Write I / O領域容量 (905)	残り容量 (906)	残り容量割合 (907)

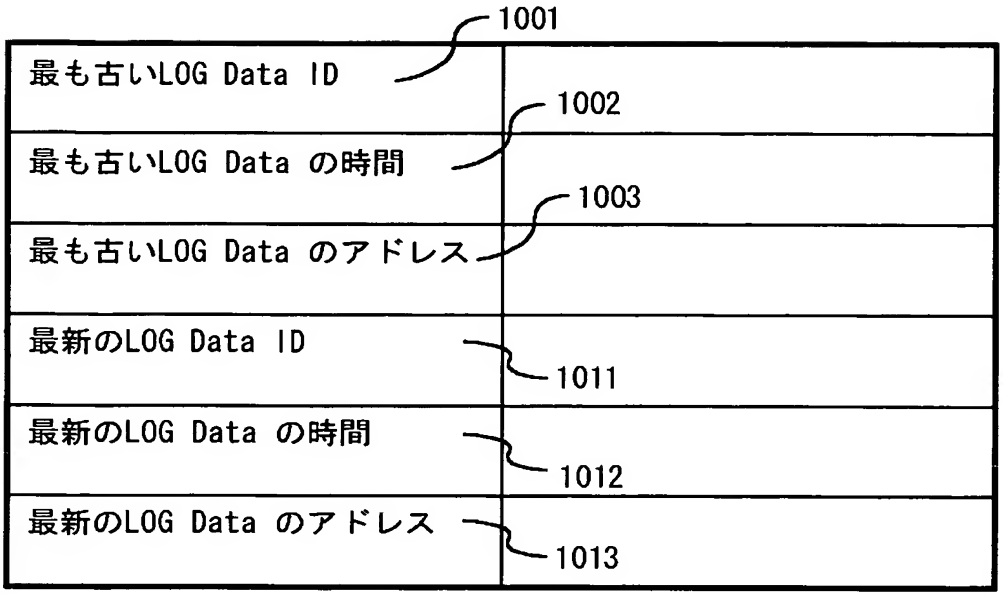


使用監視情報 (908)

【図１１】

図　１１

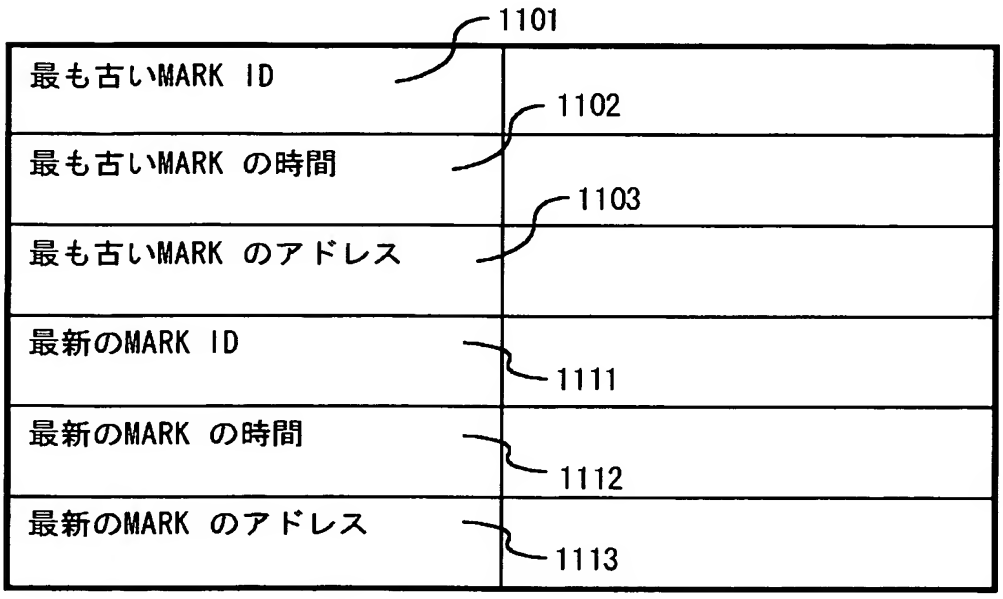
LOG ID 管理情報　157



【図１２】

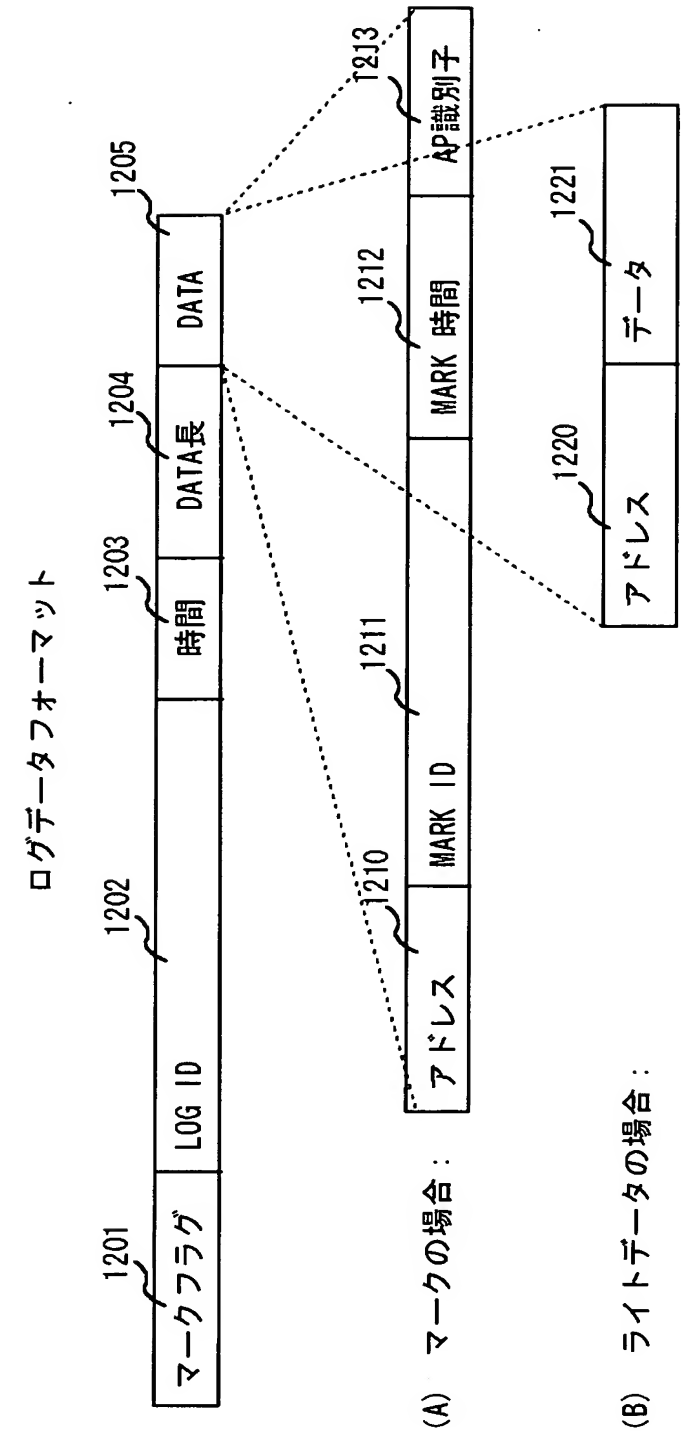
図　１２

MARK ID 管理情報　158



【図１３】

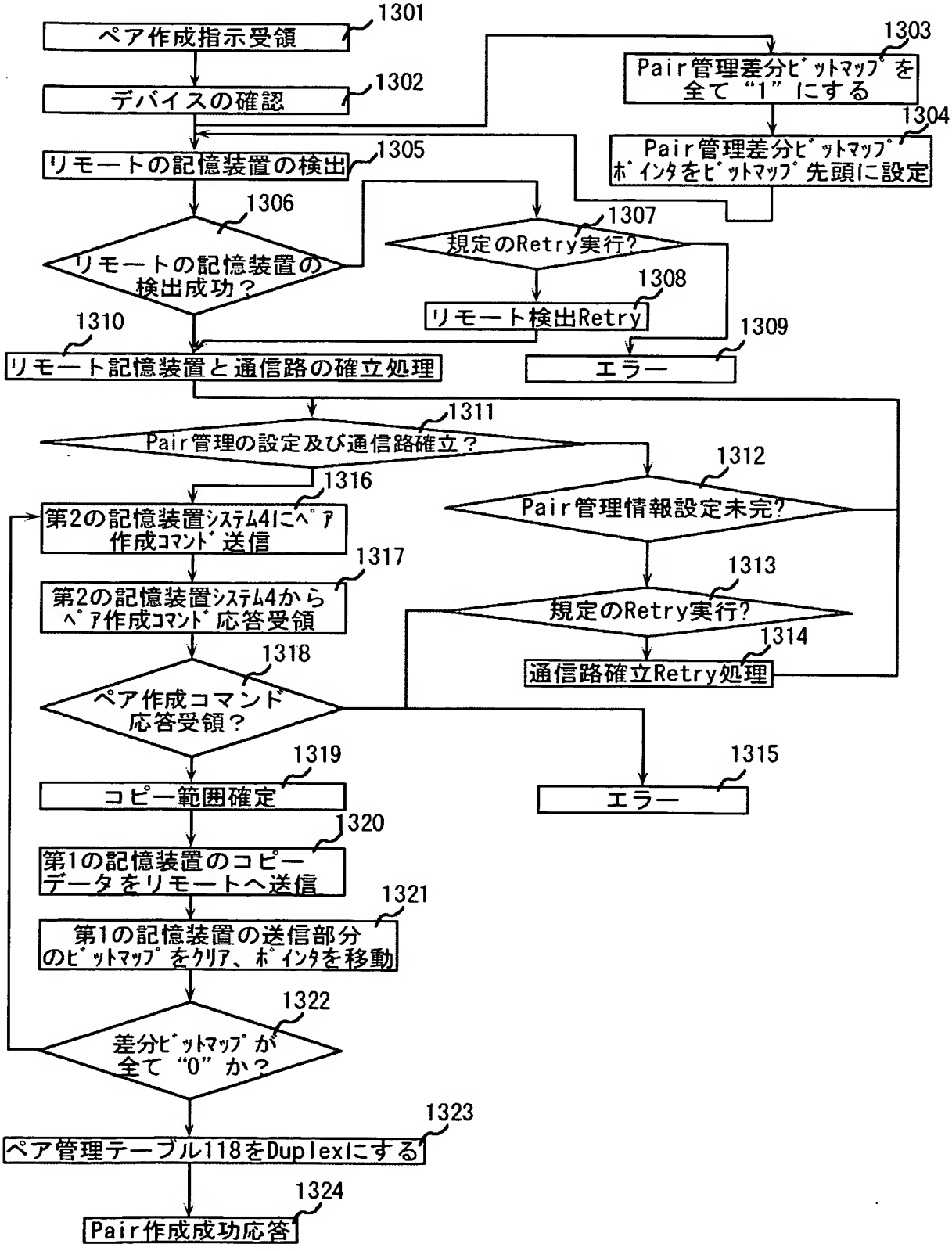
図　１３


ログデータフォーマット
1201
マーククフラグ
LOG ID
1202
時間
1203
DATA長
1204
DATA
1205
(A)　マークの場合：
1210
アドレス
MARK ID
1211
MARK 時間
1212
AP識別子
1213
(B)　ライトデータの場合：
1220
アドレス
データ
1221

【図１４】

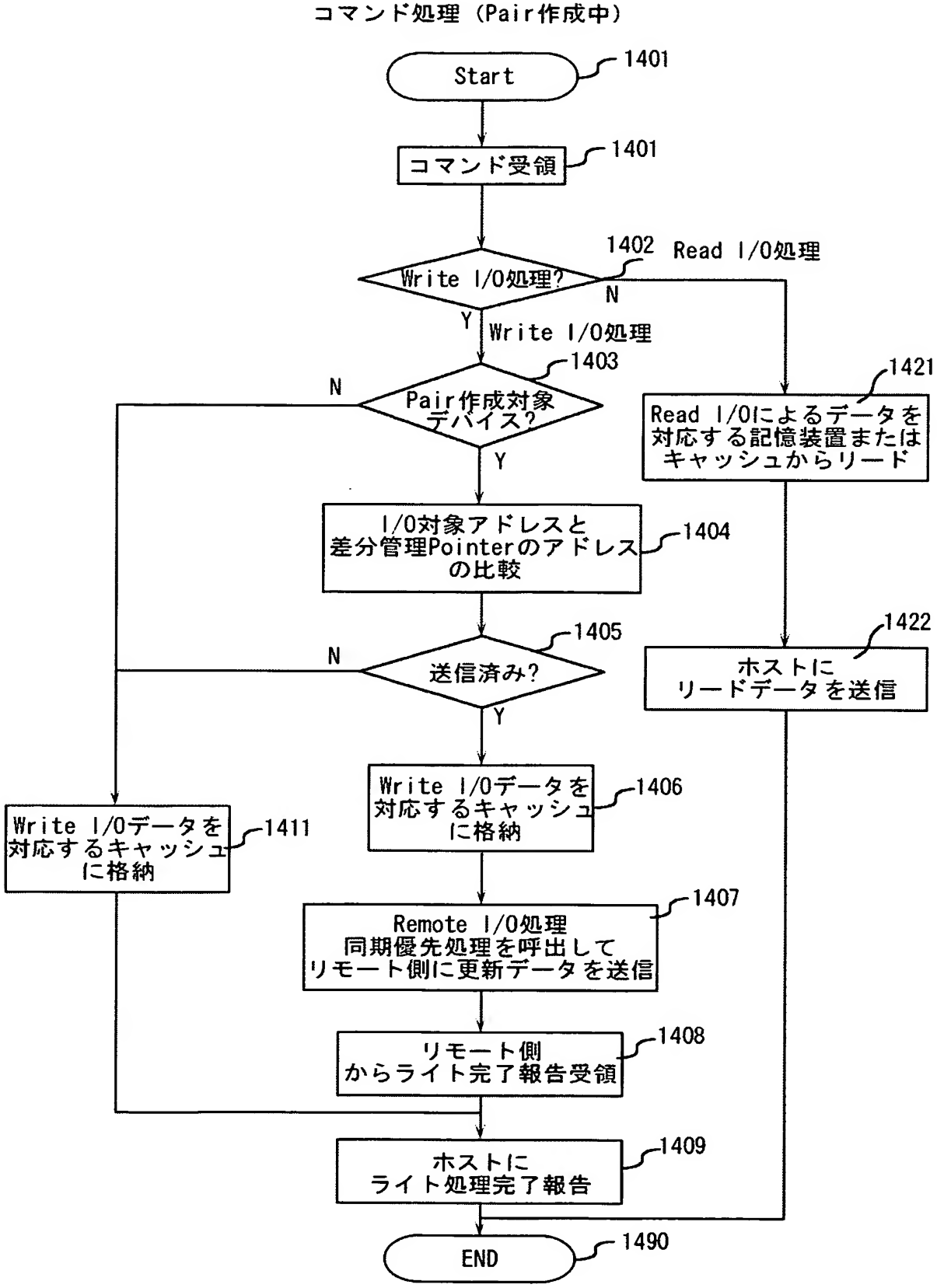
図　１４

Pair作成処理



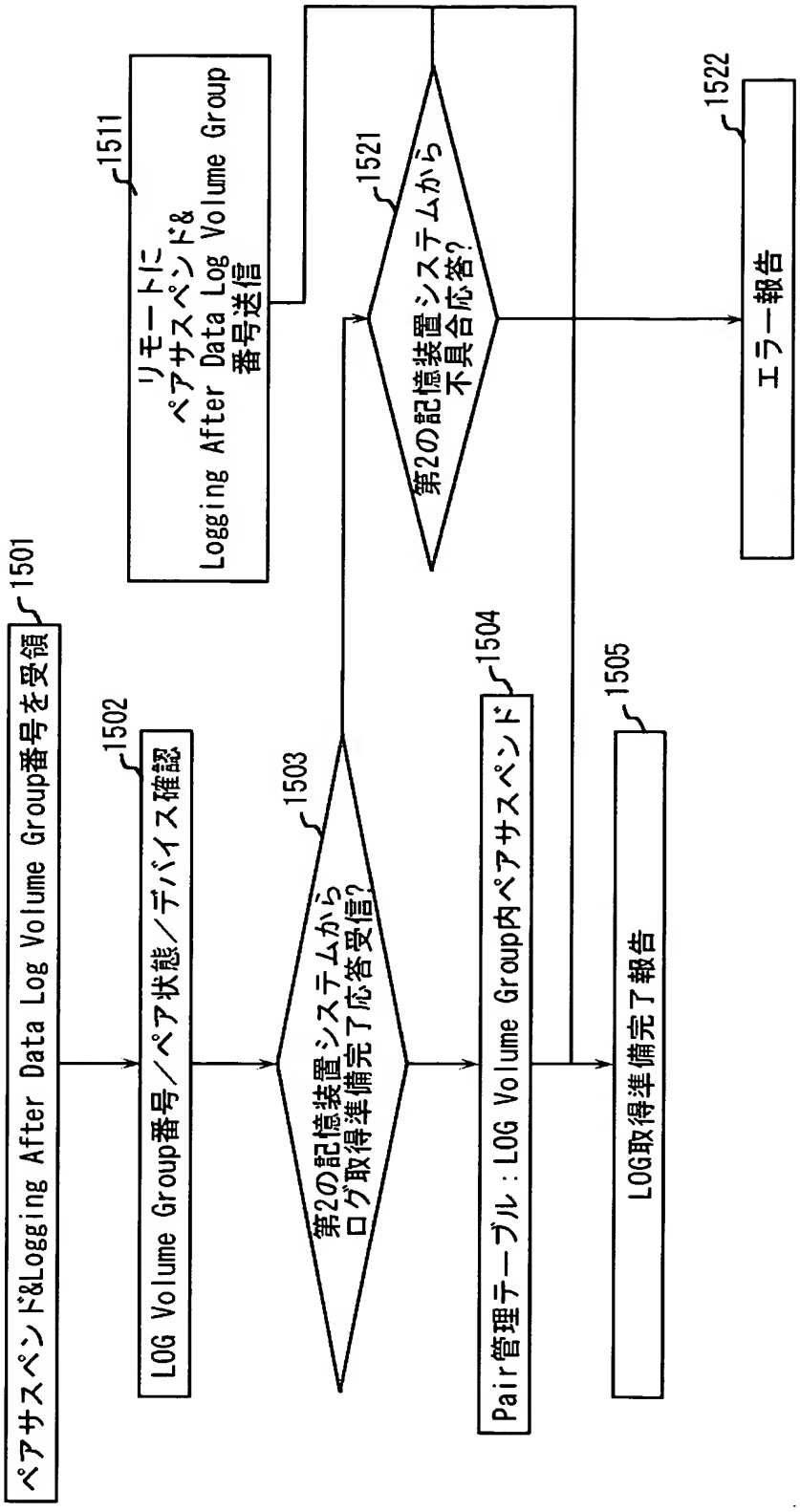
【図１５】

図　１５


コマンド処理（Pair作成中）
Start
1401
コマンド受領
1401
Write I/O処理?
1402
Read I/O処理
N
Y
Write I/O処理
1403
N
Pair作成対象デバイス?
Y
1421
Read I/Oによるデータを対応する記憶装置またはキャッシュからリード
I/O対象アドレスと差分管理Pointerのアドレスの比較
1404
1405
N
送信済み?
Y
1422
ホストにリードデータを送信
Write I/Oデータを対応するキャッシュに格納
1406
Write I/Oデータを対応するキャッシュに格納
1411
Remote I/O処理
同期優先処理を呼出してリモート側に更新データを送信
1407
リモート側からライト完了報告受領
1408
ホストにライト処理完了報告
1409
END
1490

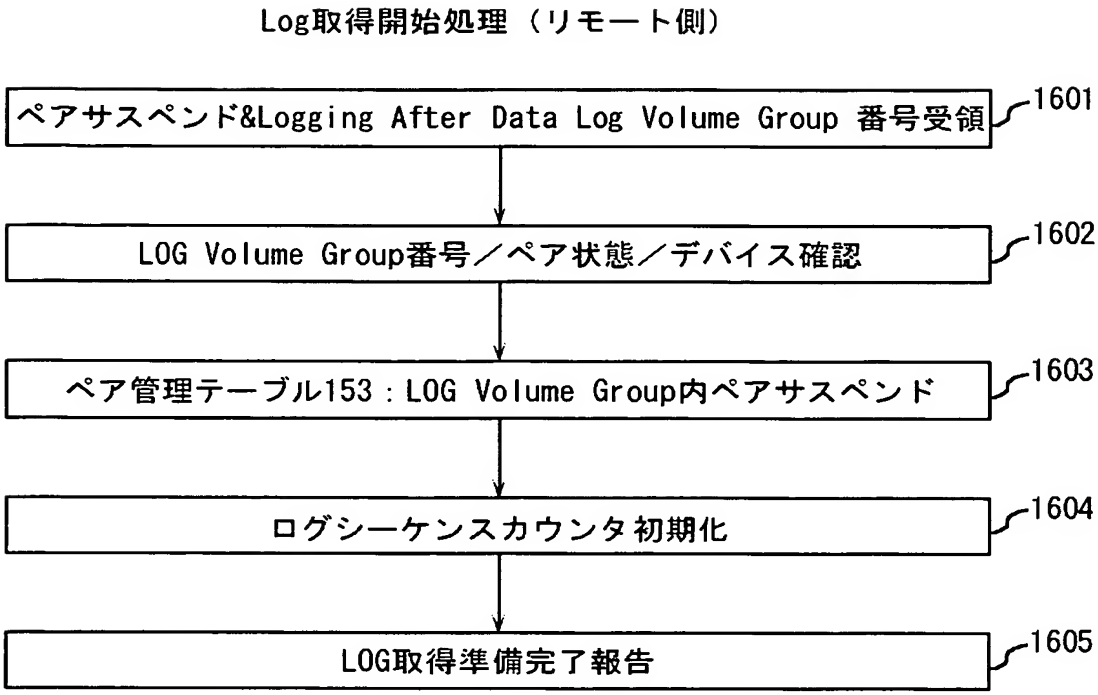
【図１６】

図　１６


Log取得開始処理（マスター側）
ペアサスペンド&Logging After Data Log Volume Group番号を受領
1501
LOG Volume Group番号／ペア状態／デバイス確認
1502
第2の記憶装置システムから
ログ取得準備完了応答受信?
1503
Pair管理テーブル：LOG Volume Group内ペアサスペンド
1504
LOG取得準備完了報告
1505
リモートに
ペアサスペンド&
Logging After Data Log Volume Group
番号送信
1511
第2の記憶装置システムから
不具合応答?
1521
エラー報告
1522

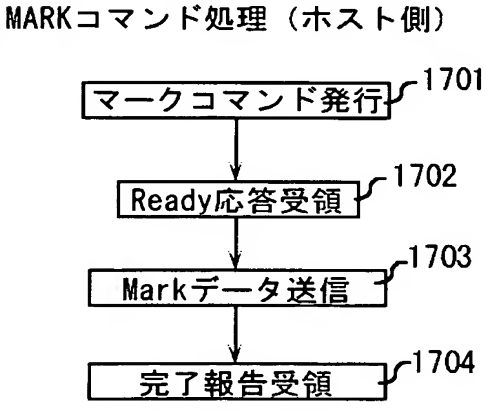
【図１７】

図　１７



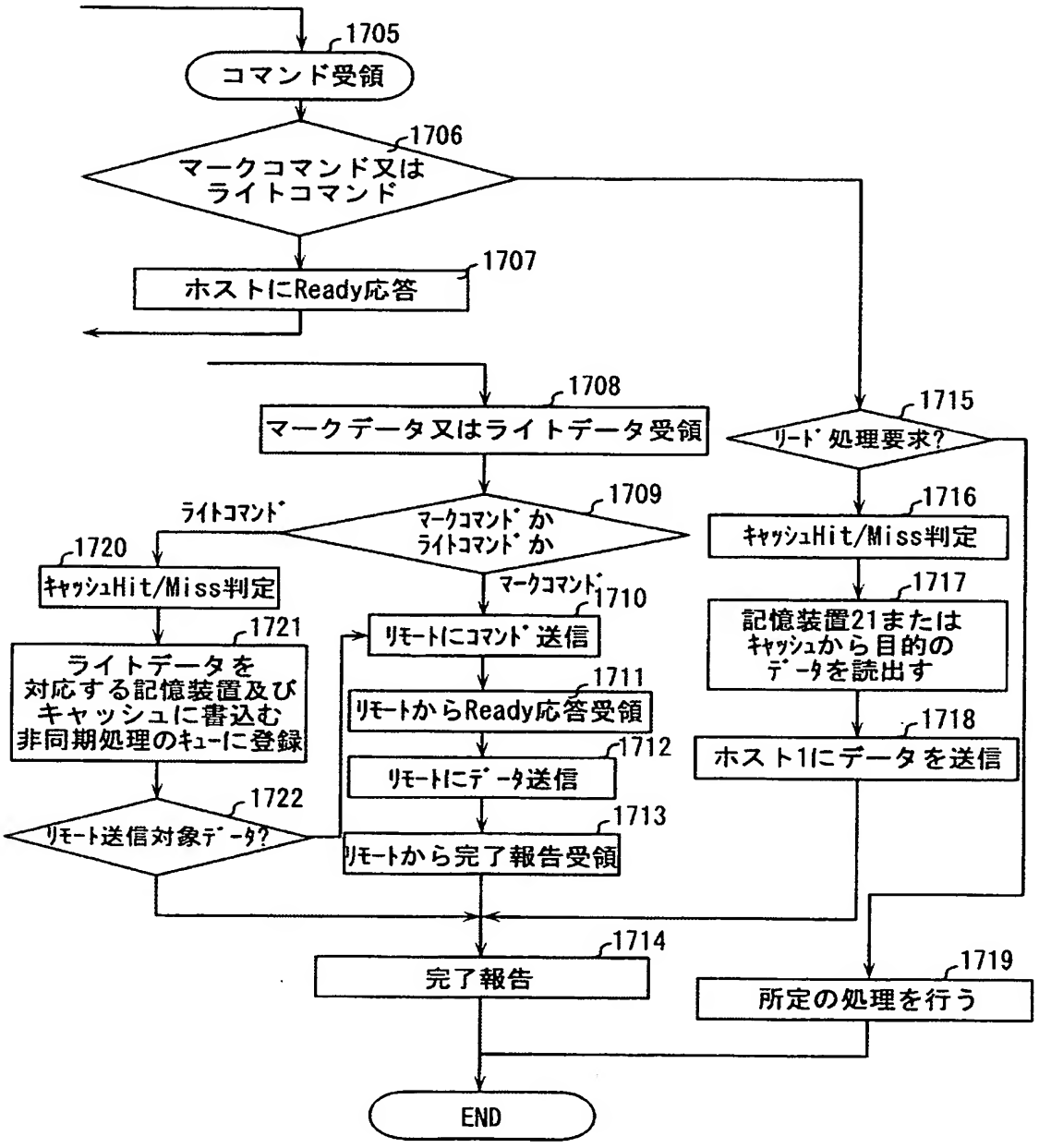
【図１８Ａ】

図　１８Ａ



【図１８Ｂ】

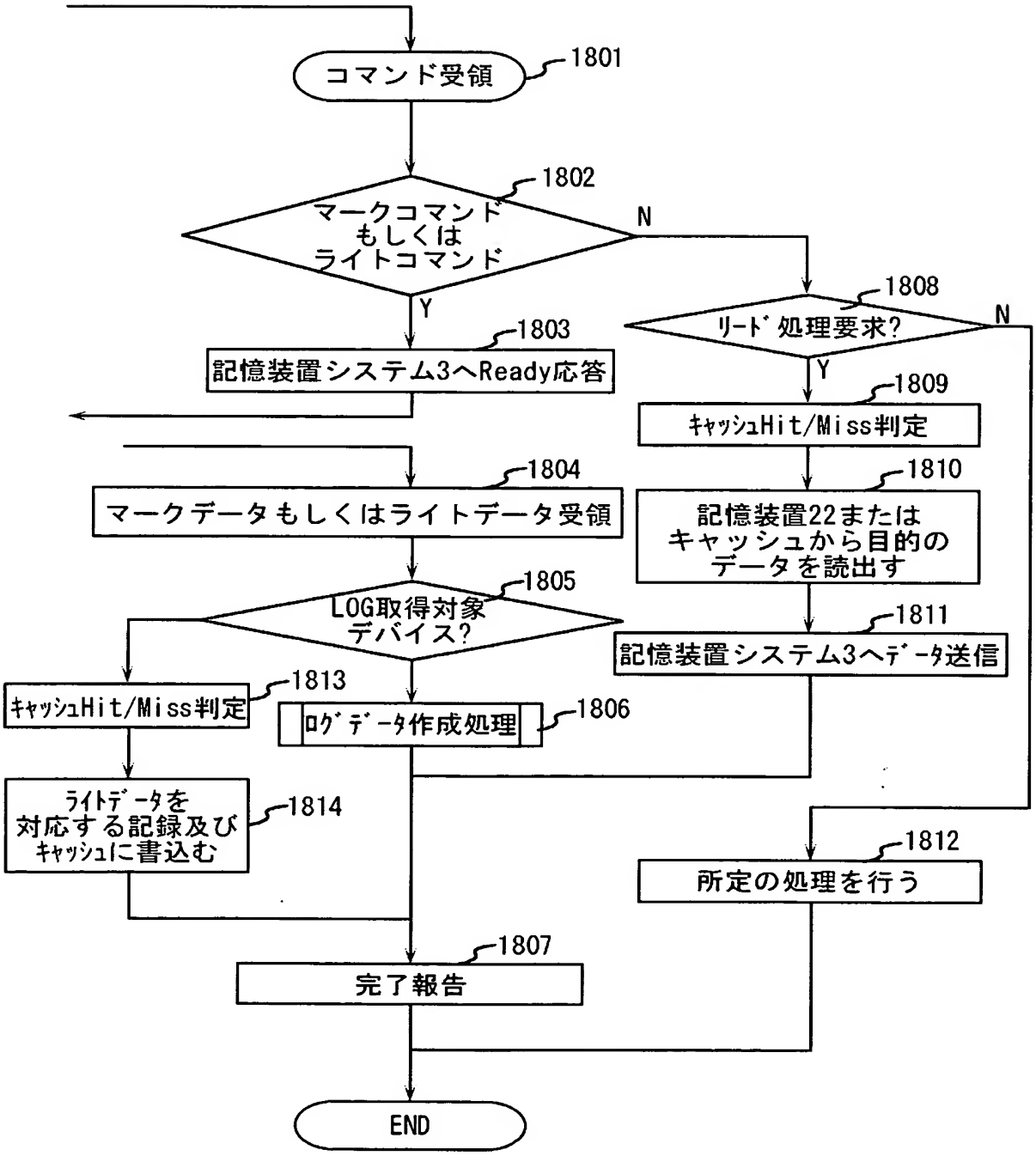
図　１８Ｂ


ＭＡＲＫコマンド処理（マスター側）
1705
コマンド受領
1706
マークコマンド又は
ライトコマンド
1707
ホストにReady応答
1708
マークデータ又はライトデータ受領
1709
ﾗｲﾄｺﾏﾝﾄﾞ
ﾏｰｸｺﾏﾝﾄﾞか
ﾗｲﾄｺﾏﾝﾄﾞか
ﾏｰｸｺﾏﾝﾄﾞ
1710
ﾘﾓｰﾄにｺﾏﾝﾄﾞ送信
1711
ﾘﾓｰﾄからReady応答受領
1712
ﾘﾓｰﾄにﾃﾞｰﾀ送信
1713
ﾘﾓｰﾄから完了報告受領
1714
完了報告
1715
ﾘｰﾄﾞ処理要求?
1716
ｷｬｯｼｭHit/Miss判定
1717
記憶装置21または
ｷｬｯｼｭから目的の
ﾃﾞｰﾀを読出す
1718
ホスト1にデータを送信
1719
所定の処理を行う
1720
ｷｬｯｼｭHit/Miss判定
1721
ライトデータを
対応する記憶装置及び
キャッシュに書込む
非同期処理のｷｭｰに登録
1722
ﾘﾓｰﾄ送信対象ﾃﾞｰﾀ?
END

【図１９】

図　１９

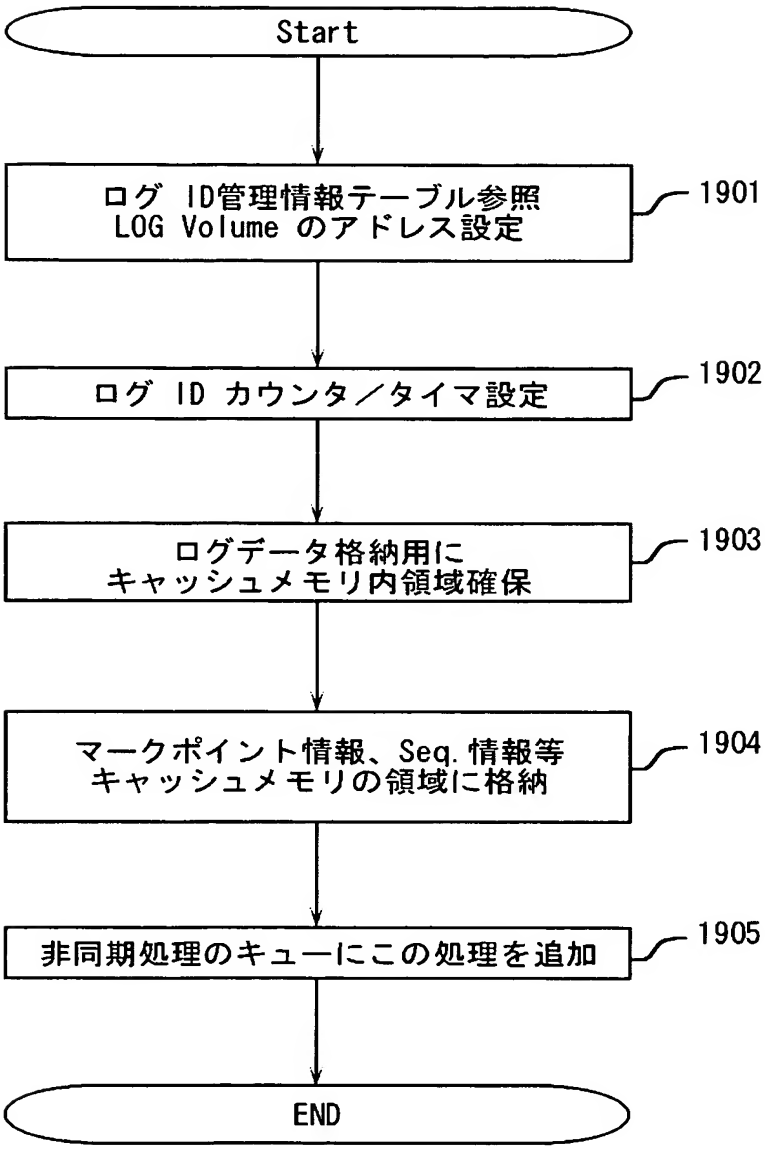
コマンド処理（リモート側）



【図２０】

図　２０

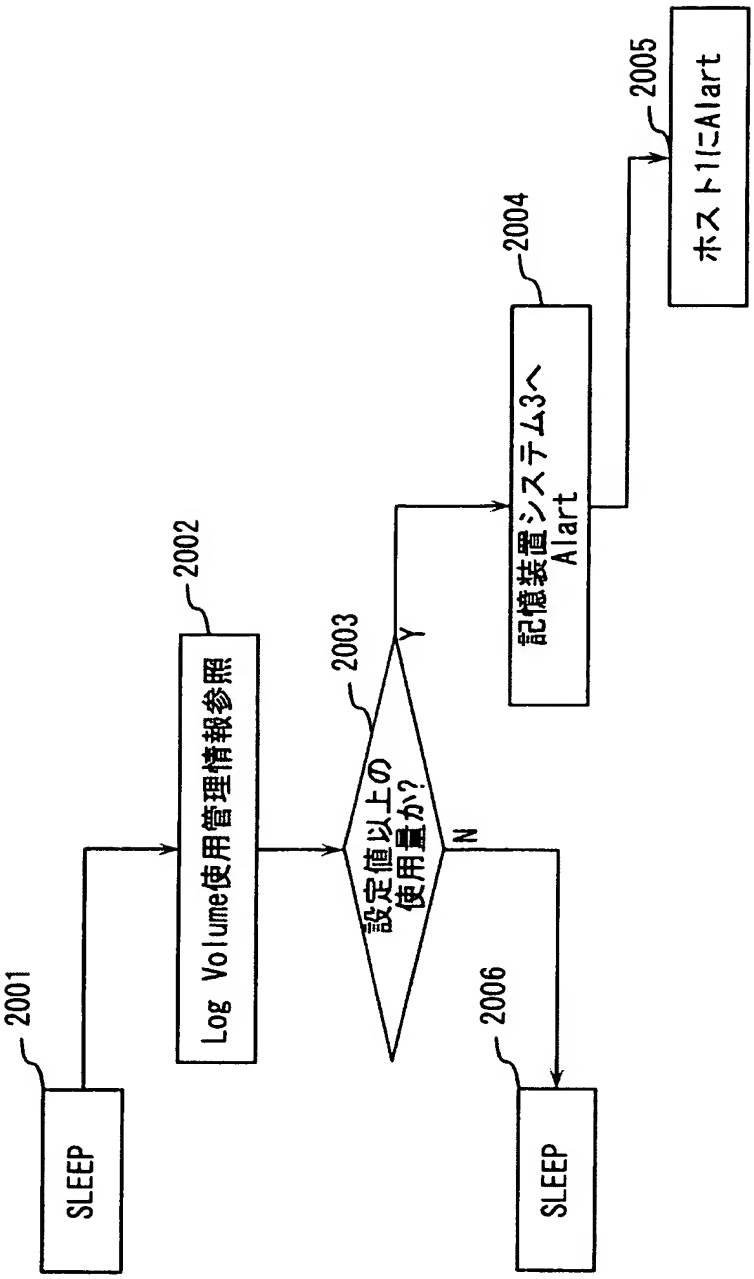
ログデータ作成処理（リモート側）



【図２１】

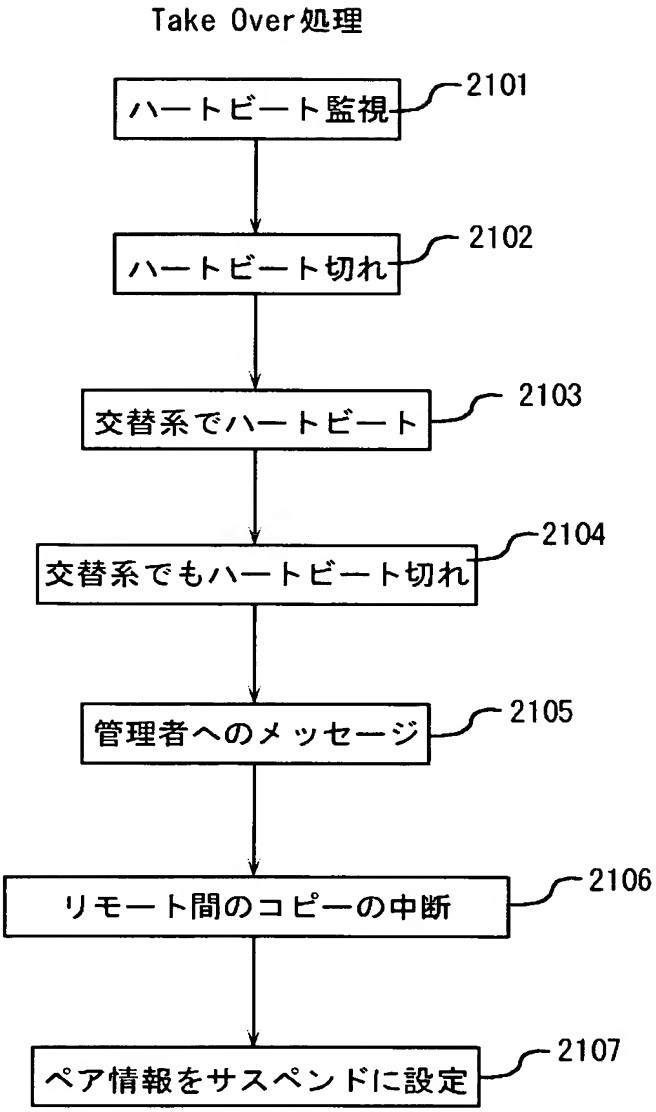
図　２１

Volume容量オーバー処理



【図２２】

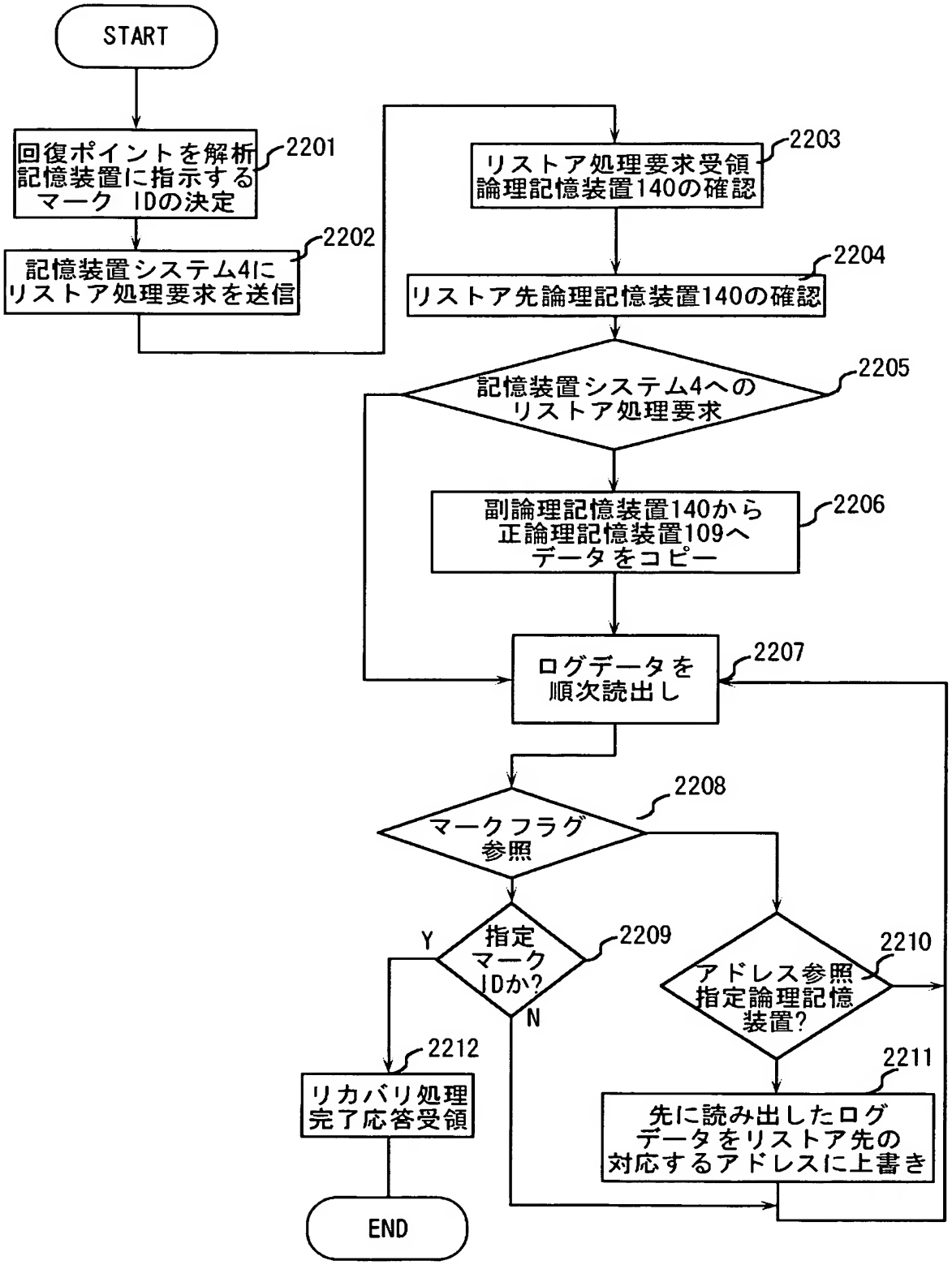
図　２２


Take Over処理
ハートビート監視
2101
ハートビート切れ
2102
交替系でハートビート
2103
交替系でもハートビート切れ
2104
管理者へのメッセージ
2105
リモート間のコピーの中断
2106
ペア情報をサスペンドに設定
2107

【図２３】

図　２３

Recovery処理



【書類名】要約書
【要約】
【課題】
ホストに負担をかけず、障害発生時点までのデータをリモートサイトで高速にリストアする。
【解決手段】
マスタ側の第１の記憶装置システムは、ホストからの入出力要求を処理し、かつリモート側の第２の記憶装置システムに対して入出力処理の結果、更新されたデータを送信し、第２の記憶装置システムは、第１の記憶装置システムからの受信したデータを更新ログデータとして保持する。ホストはアプリケーションの状態確定するコマンドをデータとして第１の記憶装置システムに送信し、第１の記憶装置システムはこのデータを第２の記憶装置システムに送信する。かつホストと第２の記憶装置システムは、コマンドに対応した識別子を双方で保持し、識別子とログデータとを関連付けることにより、ホストが任意の時点で識別子を指示することによって第２の記憶装置システムで任意の時点のデータを復元する。
【選択図】　図２

特願２００４－０２６３５６

出　願　人　履　歴　情　報

識別番号　［０００００５１０８］

１．変更年月日　１９９０年　８月３１日
［変更理由］　新規登録
住　所　東京都千代田区神田駿河台４丁目６番地
氏　名　株式会社日立製作所